PUBLIC DOCUMENTS
DEPOSITORY ITEM

NOV 26 1990

CLEMSON
LIBRARY

13. 88:PNW 263

United States
Department of
Agriculture

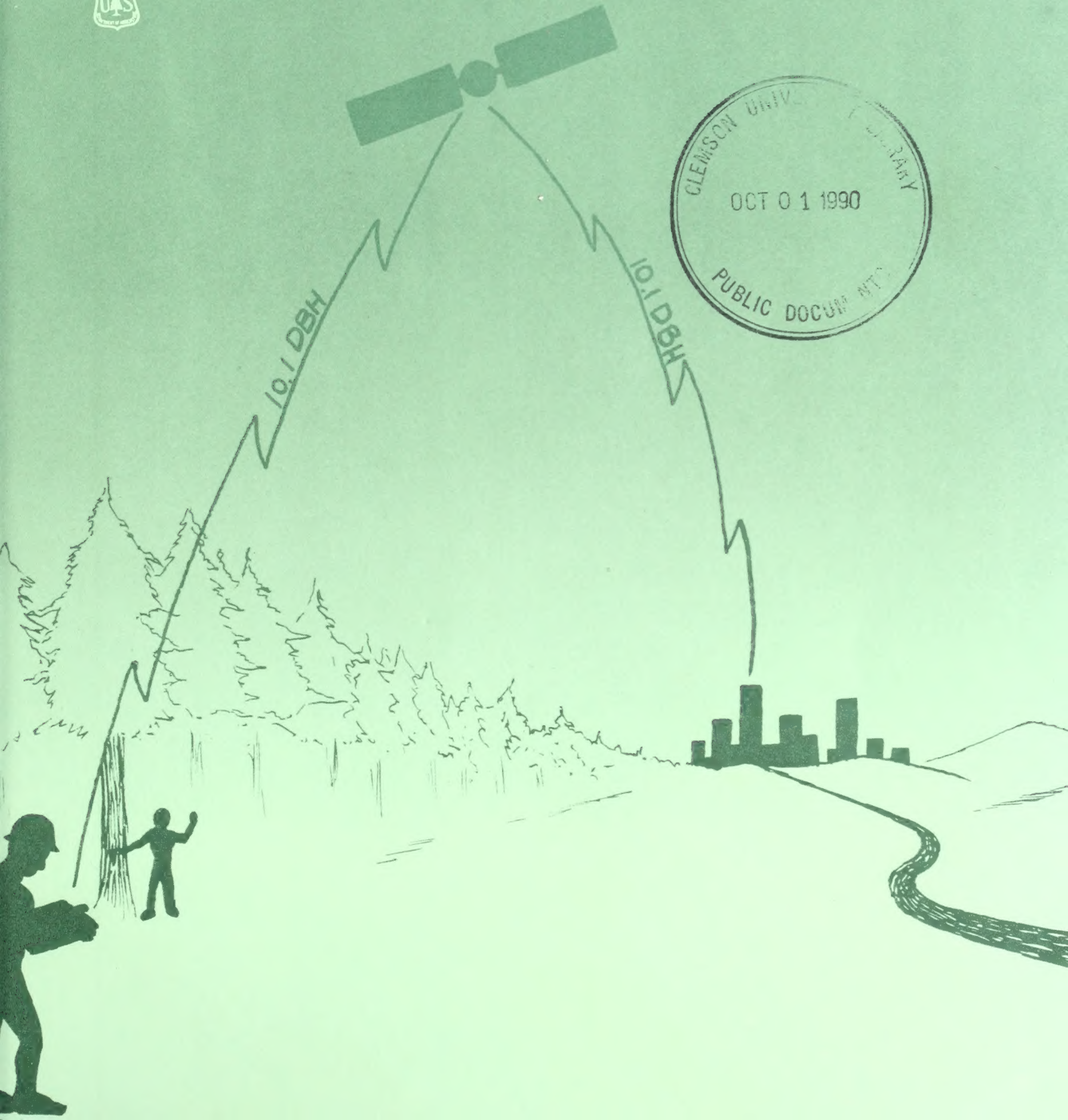
Forest Service

Pacific Northwest
Research Station

General Technical
Report
PNW-GTR-263
July 1990

State-of-the-Art Methodology of Forest Inventory: A Symposium Proceedings

CLEMSON UNIV LIBRARY
OCT 0 1 1990
PUBLIC DOCUMENTS



Digitized by the Internet Archive
in 2013

http://archive.org/details/generaltechnical263paci

STATE-OF-THE-ART METHODOLOGY OF FOREST INVENTORY: A SYMPOSIUM PROCEEDINGS

Syracuse, New York
July 30-August 5, 1989

Technical editors:
Vernon J. LaBau
Project Leader, Forest Inventory and Analysis
Pacific Northwest Research Station
Anchorage, Alaska

Tiberius Cunia
Professor of Statistics and Operations Research
College of Environmental Science and Forestry
State University of New York
Syracuse, New York

Sponsored by:
State University of New York, College of Environmental Science and Forestry
Pacific Northwest Research Station
Society of American Foresters - Forest Inventory Working Group

Cosponsors:
I.U.F.R.O. S4.02—Subject Group on Forest Resource Inventory
Society of American Foresters—Biometrics Working Group
Society of American Foresters—Remote Sensing and Photogrammetry Working Group

Published by:
U.S. Department of Agriculture, Forest Service
Pacific Northwest Research Station
Portland, Oregon
General Technical Report PNW-GTR-263
July 1990

PREFACE

The state-of-the-art of forest inventory methodology, being closely integrated with the fast-moving, high technology computer world, has been changing at a rapid pace over the past decade. Several successful conferences were held during the 1980s with the goal and purpose of staying abreast of such change.

This symposium was conceived, not just with the idea of helping make professional and technical forest mensurationists and managers aware of these changes, but also to give participants hands-on workshop experience with mensurational and inventory problems by using these latest techniques and concepts.

The objective of the symposium was to review the theory and practice of the state-of-the-art methodology of forest inventory through workshop modules, expository papers, contributed papers, posters, and discussions for those involved in teaching, research, and practice of forest inventory.

The results of the symposium, highlighted and summarized in this compendium, attest to the amount of change, the high level of interest, and the competence of the people working in forest inventory and mensuration. There are 86 contributed papers in these proceedings, many invited by the program committee, to provide a framework for the general symposium. The unique factor in this symposium was the workshops where specialists presented papers or led discussion sessions. These were followed by additional discussion periods and, in some situations, hands-on exposure to the new methodologies.

The goals and purposes of the symposium were met and, in general, exceeded. The program committee hopes that these proceedings will serve to continue to "spread the word" on the state-of-the-art of forest inventory methodology. It continues to change at an ever increasing rate, so don't take too long to contemplate this tome. We hope the reader will put some of the new ideas and concepts into practice and enjoy the thrill of using new information and techniques. Be a part of the next generation of new ideas and concepts.

PROGRAM COMMITTEE

Toby Atterbury, Atterbury Consultants, Beaverton, OR
James Barrett, University of New Hampshire, Durham, NH
Tiberius (Mike) Cunia, State University of New York, Syracuse, NY
Vernon J. (Jim) LaBau, USDA Forest Service, Anchorage, AK
Carlton Newton, University of Vermont, Burlington, VT

ACKNOWLEDGMENTS AND DISCLAIMERS

The program committee for this symposium acknowledges the support of the sponsors and cosponsors: the International Union of Forest Research Organizations (IUFRO), the Society of American Foresters, and State University of New York (SUNY), College of Environmental Science and Forestry. Special appreciation must go to SUNY for hosting the symposium. Assistance by the USDA Forest Service, Pacific Northwest Station, Portland, Oregon, in editing and publishing these proceedings is also greatly appreciated. The support of the Society of American Foresters' Inventory, Biometrics, and Remote Sensing and Photogrammetry working groups was of great importance, as was the support of the IUFRO subject group of Forest Resource Inventory (S 4.02.02).

The responsibility for technical and grammatical review and accuracy in these papers was assigned to the authors as a condition for contributing their papers. The editorial staff responsibility was limited to checks of consistency in manuscript formatting, indexing, titling, and cover preparation, as well as providing format guides and manuscript material to the authors.

The use of trade, firm, or corporation names in these proceedings is for information and convenience of the reader. Such use does not constitute official endorsement or approval by the symposium program committee or the publisher, sponsors and cosponsors of any product or service to the exclusion of others that may be suitable.

Contents

First Section: Needs, Objectives and Uses of Forest Inventories
Moderator: Charles T. Scott

The Platonic Verses and Inventory Objectives 1
H. Gyde Lund

Forest Inventory Purposes and Goals 8
Toby Atterbury

Objectives of a National Forest Inventory French NFI Meeting Users Needs 11
R.B. Chevrou

Forest Information Needs of Concerned Groups in the State of Maine 17
Don D. Tarbet and John H. Cashwell

Using Forest Inventory Data in South Australia 23
J.W. Leech

Contracting for Inventory Plot Measurement: the Minnesota Experience 29
D. Heinzen and S. Flackey

Second Section: Forest Inventory and Sampling Designs
Moderator: Timothy G. Gregoire

Line Intersect Sampling: An Historical Perspective 33
William G. Warren

Inventories of U.S. Wetlands 39
W.E. Frayer

Precise Tree Frequency Estimates in Horizontal and Vertical Point Sampling 46
Thomas W. Beers

A Comprehensive Concept for Future Forest Inventories 52
Walter Bitterlich

Efficiency of Two-Phase and Two-Stage Sampling for Tree Heights in Forest Inventory 56
H. Gary Murchison and Alan R. Ek

Sampling for Logging Residue and Waste in British Columbia 67
Stephen A.Y. Omule and David W. Ormerod

Critical Height Sampling: A Workshop on the Current State of the Technique 74
Kim Iles

Sampling and Modelling Visual Component Dynamics of Forested Areas 84
Victor A. Rudis

Model-Based Sampling 86
John C. Rennie

Importance Sampling for Volume With a Portable Computer 88
Harry T. Valentine, Timothy G. Gregoire, and George M. Furnival

Third Section: Monitoring Forests Through Succesive Inventories
Moderator: James T. Bones

An Overview of Fixed Versus Variable-Radius Plots for Successive Inventories 97
Charles T. Scott

Designing a New National Forest Inventory for Switzerland 105
Michael Koehl

Developing ATLAS Growth Parameters From Forest Inventory Plots 112
John R. Mills

Forest Change - An Assessment of Nationwide Forest Inventory Methods 119
Richard Birdsey and Hans Schreuder

Fixed-Versus Variable-Radius Plots for Change Estimation 126
Charles T. Scott and James Alegria

Assessing Pine Regeneration for the South Central United States 133
William H. McWilliams

A Permanent Plot System Designed for Making Natural Resources Management Decisions in a Rapidly Urbanizing County 141
James M. Hill, Joshua L. Royte, Jil M. Swearingen, Keith Van Ness, William Warncke, and John E. Hench

INFOR_1: An Analysis Program for Permanent Forest Inventories With Systematic Sampling 149
Piermaria Corona and Agostino Ferrara

Inventory to Support Silvicultural Decisions 158
Susan L. Stout

Funding, Pests and Hardwood Log Grades: User-Driven Cooperative Projects in the North Central States 167
M.R. Carroll

Fourth Section: On the Estimation Process In Forest Inventory
Moderator: Carl Newton

Forest Inventory: On the Structure of Error of Estimates 169
Tiberius Cunia

Use of Regression Techniques in Industrial Forest Inventories 177
Michael L. Clutter

A Multivariate Composite Estimator 181
Edwin J. Green and William E. Strawderman

The Variances of Some Compatible Estimators for the Components of Growth From Remeasured Point Samples 189
Francis A. Roesch, Jr.

The Truncated Law 197
R.B. Chevrou

Sampling Mixed Stands Using Product Estimator 204
Krishna P. Rustagi

Fifth Section: Tree Volume and Biomass
Moderator: J. Barrett

Use of Profile in Tree Volume Estimation 213
David Bruce and Timothy A. Max

Developing a Simple Stem Taper Equation From Total Stem Volume Equations or Tables 221
Piermaria Corona and Agostino Ferrara

A Sampling Procedure for Board-Foot Volume of Standing Trees 228
Timothy G. Gregoire, George M. Furnival, and Harry T. Valentine

Estimating the Volume of a Radiata Pine Stand Using Importance Sampling 237
Harry V. Wiant, Jr., and Geoffrey B. Wood

Biomass Estimates for Tropical Forests Based on Existing Inventory Data 246
Andrew J.R. Gillespie, Sandra Brown, and Ariel E. Lugo

Construction and Application of Biomass Regressions to Size-Classes of Trees 254
Xiaoming Yang and Tiberius Cunia

Fractal Geometry and Forest Measurements 260
Boris Zeide

Sixth Session: Remote Sensing and Forest Inventories
Moderator: Alan R. Ek

A Survey of Remote Sensing Methodology for Forest Inventory 267
M.J. Duggin, P.F. Hopkins, and R.H. Brock

The Use of Aerial Photographs and Angle-Gauge Sampling of Tree Crown Diameters for Forest Inventory 286
Lawrence R. Gering, Dennis M. May, and Kurt B. Teuber

Satellite Remote Sensing and Inventory of Minnesota Forest Resources 290
M.E. Bauer, T.E. Burk, A.R. Ek, S.C. Ahearn, and L.P. Queen

Inventory of the Jarrah Forest of Western Australia Using Large-Scale Photography and Ground Sampling 298
Paul Biggs

A Practical Example of Forest Mapping and Timber Volume Estimation Using Landsat-5 TM Imagery 303
David Chung Liang Lee, René Antonio Novaes, Pedro Hernandez Filho, and Flávio Jorge Ponzoni

An Aerial Photographic Key as an Aid to Improving Photointerpretation Accuracy 310
Rachel Riemann Hershey

Photointerpretation Accuracy Across Two Decades of Forest Inventory 318
G. Deegan and W.A. Befort

On Using Non-Imageable Point for Measuring Object Height on Vertical Aerial Photograph 324
C.J. Liu

Organizing Statewide 35mm Photography in Support of Forest Inventory 332
W.A. Befort and G. Deegan

A Study of Pinus Mugo-Stands in the Karwendel Mountain Region (Tyrol, Austria) 336
Josef Hackl and Erwin Knappitsch

Seventh Section: Forest Health and Decline
Moderator: T. Cunia

Environmental Health Concerns: A Role for Forest Inventory and Monitoring 341
Joseph E. Barnard

Monitoring the Health of White Pine in New Hampshire (1959-1988) 349
John P. Jastrzembski and James P. Barrett

National Inventories and Inventories of Endangered Forests in Europe 356
Michael Koehl

Indicator Plants in Forest Health Surveys 364
Paul C. Van Deusen and Joseph E. Barnard

Eighth Section: Spacial Analysis of Inventory Data
Moderator: Joseph C. Mawson

Spatial Analysis of Forest Inventory Data 369
Carlton M. Newton and Steven T. Bower

Spatial Analysis Using a Systematic Grid of Forest Inventory Plots: Selected Examples for the South Central United States 379
William H. McWilliams and Joanne L. Faulkner

Application of Permutation Techniques in Analyzing Spatial Patterns 384
Robin M. Reich, Paul W. Mielke, Jr., and Frank G. Hawksworth

Ninth Section: Forest Inventory and Geographic Information Systems
Moderator: Toby Atterbury

Geographic Information Systems: Where Do We Stand? 391
Lee Pierce Herrington

Growth Models and Geographic Information Systems 397
Sarah Osborne

Building Predictive Models from Geographic Information Systems 403
David L. Verbyla

Tenth Section: Mathematical Models and Forest Inventory
Moderator: John C. Rennie

Status and Future of Growth and Yield Models 409
Harold E. Burkhart and Thomas M. Brooks

Interfacing Forest Inventory With Growth and Yield Models 415
Alan R. Ek

Modeling With Data of the First Swiss National Forest Inventory 420
Edgar Kaufmann

Workshop 1: Updating Methods for Forest Inventories
Moderator: Richard A. Birdsey

Updating Methods for Forest Inventories - An Overview 429
Richard Birdsey

Remote Sensing Techniques for Updating Forest Inventories 436
Kurt B. Teuber

Inventory Updates: Overview of the North Central Experience 442
W. Brad Smith

Updating Inventories in Western Oregon and Washington Using Tree Growth, Mortality and Harvest Models Developed from Permanent Plot Data 449
Colin D. MacLean and Neil McKay

Kalman Filter to Update Forest Cover Estimates 457
Raymond L. Czaplewski

Workshop 2: Integrating Multiple Value Forest Surveys Into Timber Survey
Moderator: Victor A. Rudis

The Importance of Forest Inventories to the Conservation of Biological Diversity 467
Richard E. Rice

The Forest Inventory: Incorporating Fisheries and Wildlife Values 475
Frances A. Hunt

Linking Multiple Resources Analyses to Land Use and Timber Management: Application and Error Considerations 478
Curtis H. Flather, Linda A. Joyce, and Rudy M. King

Using Land Base Inventories to Access Range and Wildlife Resources Nationally 486
Linda A. Joyce and Curtis H. Flather

Multiple Value Forest Surveys in the Midsouth States 495
Victor A. Rudis

Expanding Conventional Timber Inventories into a Broader, Multiresource Inventory in the Southeastern States 505
Noel D. Cost

Developing Integrated Forest Surveys for the Rocky Mountain States 512
Renee O'Brien

Nonindustrial Wood Resources in the Rocky Mountain States 517
J. David Born

Multiresource Inventories in Alaska 522
Bert Mead and Vernon J. LaBau

Multiresource Inventories in the Pacific Coast States—Progress and Future Direction 529
Janet L. Ohmann

State-of-the-Art Methodology of Forest Inventory: Wildlife Habitat Assessment in the Northeastern United States 538
Robert T. Brooks

Workshop 3: Accessing USDA Forest Service Forest Inventory and Analysis (FIA) Data
Moderator: Mark Hansen

The Eastwide Forest Inventory Database 545
Mark H. Hansen

Access to Forest Inventory Data for the Southeast 553
Raymond M. Sheffield

Interactive Access to Forest Inventory Data for the South Central United States 556
William H. McWilliams

North Central Forest Inventory Data on a Microcomputer 558
Jerold T. Hahn and Mark H. Hansen

Workshop 4: Surveying With Partial Replacement
Moderator: T. Cunia

Successive Forest Inventories: Basic SPR Approach for Two Measurements 565
Tiberius Cunia

Successive Forest Inventories: Extensions of the Basic Approach 572
Tiberius Cunia

Successive Forest Inventories: Error of Biomass Regressions Part I: CFI With Permanent Plots Only 580
Tiberius Cunia

Successive Forest Inventories: Error of Biomass Regressions Part II: CFI With Partial Replacement of Plots 586
Tiberius Cunia

Author

Ahearn, S. C.	290
Alegria, James	126
Atterbury, Toby	8
Barnard, Joseph E.	341, 364
Barrett, James P.	349
Bauer, M. E.	290
Beers, Thomas W.	46
Befort, W. A.	318, 332
Biggs, Paul	298
Birdsey, Richard	119, 429
Bitterlich, Walter	52
Born, David	517
Bower, Steven T.	369
Brock, R. H.	267
Brooks, Robert T.	538
Brooks, Thomas M.	409
Brown, Sandra	246
Bruce, David	213
Burk, T. E.	290
Burkhart, Harold E.	409
Carroll, M. R.	167
Cashwell, John H.	17
Chevrou, R. B.	11, 197
Clutter, Michael L.	177
Corona, Piermaria	149, 221
Cost, Noel D.	505
Cunia, Tiberius	169, 254, 565, 572, 580, 586
Czaplewski, Raymond L.	457
Deegan, G.	318, 332
Duggin, M. J.	267
Ek, Alan R.	56, 290, 415
Faulkner, Joanne L.	379
Ferrara, Agostino	149, 221
Filho, Pedro Hernandez	303
Flackey, S.	29
Flather, Curtis H.	478, 486
Frayer, W. E.	39
Furnival George M.	88, 228
Gering, Lawrence R.	286
Gillespie, Andrew J. R.	246
Green, Edwin J.	181
Gregoire, Timothy G.	88, 228
Hackl, Josef	336
Hahn, Jerold T.	558
Hansen, Mark H.	545, 558
Hawksworth, Frank G.	384
Heinzen, D.	29
Hench, John E.	141
Herrington, Lee Pierce	391
Hershey, Rachel Riemann	310

Hill, James M.	141
Hopkins, P. F.	267
Hunt, Frances A.	475
Jastrzembski, John P.	349
Joyce, Linda A.	478, 486
Iles, Kim	74
Kaufmann, Edgar	420
King, Rudy M.	478
Knappitsch, Erwin	336
Koehl, Michael	105, 356
LaBau, Vernon J.	522
Lee, David Chung Liang	303
Leech, J. W.	23
Liu, C. J.	324
Lugo, Ariel E.	246
Lund, H. Gyde	1
MacLean, Colin D.	449
McKay, Neil	449
Max, Timothy A.	213
May, Dennis M.	286
McWilliams, William H.	133, 379, 556
Mead, Bert	522
Mielke, Paul W., Jr.	384
Mills, John R.	112
Murchison, H. Gary	56
Newton, Carlton M.	369
Novaes, René Antonio	303
O'Brien, Renee	512
Ohmann, Janet L.	529
Omule, Stephen A. Y.	67
Ormerod, David W.	67
Osborne, Sarah	397
Ponzoni, Flàvio Jorge	303
Queen, L. P.	290
Reich, Robin M.	384
Rennie, John C.	86
Rice, Richard E.	467
Roesch, Francis A., Jr.	189
Royte, Joshua L.	141
Rudis, Victor A.	84, 495
Rustagi, Krishna P.	204
Schreuder, Hans	119
Scott, Charles T.	97, 126
Sheffield, Raymond M.	553
Smith, W. Brad	442
Stout, Susan L.	158
Strawderman, William E.	181
Swearingen, Jil M.	141
Tarbet, Don D.	17
Teuber, Kurt B.	286, 436

Valentine, Harry T. 88, 228
Van Deusen, Paul C. 364
Van Ness, Keith 141
Verbyla, David L. 403
Warncke, William 141
Warren, William G. 33
Wiant, Harry V., Jr. 237
Wood, Geoffrey B. 237
Yang, Xiaominng 254
Zeide, Boris 260

Needs, Objectives and Uses of Forest Inventories
Moderator: Charles T. Scott

H. Gyde Lund, The Platonic Verses and Inventory Objectives

Toby Atterbury, Forest Inventory Purposes and Goals

R.B. Chevrou, Objectives of a National Forest Inventory French NFI Meeting Users Needs

Don D. Tarbet and John H. Cashwell, Forest Information Needs of Concerned Groups in the State of Maine

J.W. Leech, Using Forest Inventory Data in South Australia

D. Heinzen and S. Flackey, Contracting for Inventory Plot Measurement: the Minnesota Experience

The Platonic Verses and Inventory Objectives

H. Gyde Lund
Inventory Forester
USDA Forest Service, Timber Management Staff
P.O. Box 96090
Washington, DC 20090-6090

ABSTRACT

Platonic verses *are writings that tend toward purely spiritual or ideal situations such as those found in the Bible and other sacred writings or in some of our natural resource management laws and national direction.* ***Inventory objectives*** *are summary statements of the desired end results of surveys of goods and materials in stock.* The process of converting information requests to inventory objectives is discussed and personal thoughts on additional inventory objectives, use of existing information, subjective sampling, moving plot locations, and use of allowable sampling errors are given.

In the Beginning...

An inventory is an accounting of goods and materials in stock. The following steps are usually followed to implement an inventory:

1) Define the objectives for the inventory;

2) Determine the information needs, standards, and methods to collect the data to meet objectives;

3) Gather the data; and

4) Compute and report the results.

Of these, step number one is often the most difficult to achieve. This is because objectives for inventory often are not spelled out clearly. Consider the following direction extracted from the Forest and Rangeland Renewable Resource Planning Act of 1974 (USDA Forest Service, 1983):

> Sec. 3. Renewable Resource Assessment. *The Assessment shall...include but not limited to:... (2) An inventory, based on information developed by the Forest Service and other Federal agencies, of present and potential renewable resources....*
>
> Sec. 5. National Forest Systems Resource Inventories. *As part of the Assessment, the Secretary of Agriculture shall develop and maintain on a continuing basis a comprehensive and appropriately detailed inventory of all National Forest Systems lands and renewable resources. This inventory shall be kept current so as to reflect changes in conditions and identify new and emerging resource values.*
>
> Sec. 12. Renewable Resources. *In carrying out this Act, the Secretary of Agriculture shall utilize information and data available from other Federal, State, and private organizations and shall avoid duplication and overlap of resource assessment and program efforts of other Federal Agencies.*

The direction is clear that the USDA Forest Service is to inventory some resources and to avoid duplication of effort, but based on the information contained in the laws, we would be hard-pressed to design an inventory to ensure that it would meet the needs that Congress had in mind.

How then do we determine inventory objectives from these and other such platonic verses and what additional things do we need to consider in designing inventories? These topics are the subjects of this paper.

The Book of Genesis

And on the eighth day, the LORD God told man to go forth and inventory the trees and herbs of the land, the beasts of the field, the fish of the seas, and the fowl of the air so that their numbers be knownst (Genesis 2:8(a)).

Given no funding, access, and personnel, Adam and Eve were faced with a very formidable task. They tried to take God's command literally. They failed and were cast out of the Garden of Eden.

Over the next few millenniums man, plants, and animals continued to multiply. In spite of increased person-power, God still did not have a suitable inventory, so he sent a great flood to reduce the number of life-forms to "two of every living thing". When the floods receded God knew what he had.

But life prospered and was fruitful and multiplied. And it came to pass that once again God was faced with an inventory problem. Even such acts as turning people into pillars of salt did not bring about an adequate survey of the earth's resources.

Frustrated, God delegated his inventory needs to the Pharaohs. Soon, Joseph was hired as an advisor to the rulers. Capitalizing on a clause in Joseph's contract under "other duties as assigned", the Pharaohs soon got on his case to get an inventory done. Joseph realized that he needed more information from the powers that be in order to do a suitable job. So he asked the Pharaohs:

1. *What decisions are going to be made on the basis of the inventory?*

2. *What information is really needed to make these decisions?*

3. *What is the area for which the estimates are to apply?*

4. *When are the results needed and in what form?*

5. *What are the constraints on funding and people power to get an inventory done?*

Joseph also realized that man could not inventory the resources by complete enumeration in every instance. In a dream, the theory of sampling came to him. And it came to pass, that to implement the sampling theory, Joseph developed some additional questions to be asked of the Pharaohs. These included:

1. *What impact will information errors have on the decisions being made.*

2. *What impact will incorrect decisions have on the resources?*

3. *What impact will incorrect decisions have on the decision-maker?*

4. *And most importantly, what impact will incorrect decisions have on the inventory designer?*

To properly design an inventory, the inventory designer must have an understanding of the importance of the results, how they will be used, and act accordingly. In order to have this understanding, the designer must communicate with the person requesting the information. Answers to the questions given above and some knowledge about the requestor will help develop the inventory objectives. Atterbury (1987) goes into considerable detail about the types of questions to be asked for specific uses.

The effort devoted to questioning will pay off in the long run but the amount of effort expended depends on the source of the request. If a request comes from Miss Sarah Jane in Ms Bixby's fourth grade class in Tulsa, Oklahoma for a science project she was working on, we would provide the information requested as best we can. We would not necessarily launch a multi-million dollar project to answer her question if the data were not available.

If, on the other hand, requests come from Congressional or White House aides, we work with them to define their needs in terms of the kinds of data we have available. From this we can judge if we can provide the information or not. If we can't, then we may have to lay out some options for gathering the data; balancing the time when the information is needed against the person-power and funding available.

The Book of Numbers

Take ye the sum of all the trees of the forest, by the nation in which they stand with thy names and uses from two-tenths cubits at breast height and upwards (Numbers 1:2(a)).

By the time Moses floated into the scene, the Pharoahs were speaking in terms that the inventory designer could understand. Moses realized that one could have different inventories to meet different needs or multiple objectives from the same inventory. For example, in the USDA Forest Service, we have information needs for policy and program development, land use planning, and activity planning. Each of these needs could be served by one or several data collection efforts.

Even though one has clearly defined the objective statement, one must also be aware of the additional or hidden needs.

As I prepared this paper, I queried our Forest Service Regions and Forest Inventory and Analysis (FIA) Staffs to find out what the field units saw as the objectives for their inventories. Nearly all units that responded, quoted or referred to the platonic statements found in our laws and national direction. Yet, while these may be the primary reasons for the inventory, there are reasons, that could lead one to question why we use the designs that we do.

For example, the primary objective of the Forest Inventory and Analysis units is to provide data for National Assessments and State Survey Reports. A secondary objective is to maintain a net-work of permanent plots for monitoring and research activities. This objective has more influence on the inventory design than does the primary objective.

A similar statement may be made of National Forest System (NFS) inventories. The primary objective is to provide data for Forest planning and National Assessments. A hidden objective is to use the inventories for project planning. By recognizing all needs from the start, it will make the design and understanding of inventories much more simpler.

Lund (1986a) outlines some of the inventory needs we have in the federal government that are over and above those stated in the laws.

CLEMSON UNIVERSITY LIBRARY
PUBLIC DOCUMENTS DEPT.

The Book of Exodus

And it came to pass, that the Pharaohs turned over their inventory duties to the data collectors. The COLLECTORS spake all these words saying,... (Exodus 20:1(a)).

In recent years, four inventory commandments appear to have evolved that merit some discussion.

1. ***Thou shalt not use any work other than thine own*** (Exodus 20:3(e)).

For some reason, we (myself included) often ignore the work done by others. Every time a new inventory specialist is hired, out goes the old survey and in comes a new one. This is indeed unfortunate for several reasons. First the continuity is lost between inventories. Results are not easily compared because of the differences in designs. Secondly, a great deal of valuable data and knowledge is ignored creating a loss of a previous investment.

There is a vast amount of existing information available in the United States from a variety of sources. Yet we often do not take the time to search for it or if we find it we do not use it because we did not collect it or it was not to our standards.

As I work with National Forests and Forest Inventory and Analysis Units I have come to realize that there are people out there at least as smart as I am, and that their information is of considerable value. Before designing a new inventory, make sure that the information does not exist or that it is not adequate. Guidance on evaluating existing information is provided in Lund (1986b).

There is not only a problem with using other people's data, but there is a problem of sharing data we collect ourselves. Some of this may be attributed to problems with definitions and standards, but most is due to three things:

a. *Fear of loosing control.* Knowledge is power, and power is difficult to share. If we share our information or knowledge, we may no longer be looked upon as the "authority".

b. *Fear of someone using the data against us.* This is especially true where resource management decisions are frequently challenged in the courts.

c. *Fear of someone scooping our research.* We have not as yet published, therefore, if we released the data our competitors may beat us to the punch.

While these fears may be valid, at least for most federal employees, we are mandated to share our information under the Freedom of Information Act when requested.

2. ***Thou shalt not use subjective sampling*** (Exodus 20:14(f)).

Not all functions use random sampling in their inventories. In our national direction in the Forest Service, we now require or inventories to be "scientifically valid" instead of "statistically valid". Nearly everything we do naturally, from the selection of mates to selection of food, shelter and clothing, jobs, etc. is based on deliberate sampling. Why then when we want to inventory our resources, do we use random sampling?

The basic arguments are that random sampling eliminates personal biases and creates objectivity so that the reliability of the inventory can be established. The disadvantages are the cost and time in creating the sampling frame and in building in the random factor.

Arguments for deliberate sampling (which includes purposive sampling, subjective sampling and I suppose the emerging techniques of model-based sampling and importance sampling) are that it is sometimes, if not most often, cheaper and logical in the mind of the user.

Subjective sampling may not as biased as we are lead to believe. Assume for example, we wanted to select a person who represented the average weight of people in this room. All of you would agree, that I would should not be selected (in a random draw, I could be chosen). Others in the room may be eliminated from the sample. Arguments would occur closer to the actual mean rather than at the extremes. Consequently the margin that a subjective sample would vary would be relatively narrow. The costs of creating randomness and measuring things that clearly are not representative are avoided and the cruiser can get on with the job sooner.

In the past, I have only recommended subjective sampling when: the person that collects the data will be the only one that will ever use them; or, when it is used as a means to get an estimate of the variation for determining sampling intensity for a bona fide statistical sample. I still believe that these two uses are valid. But there are other circumstances where subjective sampling is used and the results may be logically more valid than a statistical sample.

Assume we have a ten-thousand acre forest that has had two inventories. One inventory was based on a randomly located grid of 20 plots. In the other inventory, the entire forest was divided and mapped into 200 stands ranging in size from 10 to 100 acres each. Within each stand, two plots were established at locations that the cruiser determined to be representative of the stand itself. In both inventories the same plot design was used, the same data were collected, and the same statistics were generated. Naturally the results differed.

If you had to invest money based on the results of the inventory, which inventory would you accept as being the most authoritative? I would side with the more intense subjective sample. The reason for this is simple -- there is a much higher intensity of samples distributed across the population of interest. Certainly the stand-based inventory could be improved from a statistical view point by randomly selecting plots within each stand and the statistical survey could be strengthened by adding more plots, but inventories should not be summarily rejected simply because subjective sampling was involved.

The use of deliberate sampling depends again on the objectives of the inventory and the intended audience. The more clearly objectives of the survey and the audience are defined, the more clearly one can design the inventory.

3. ***Thou shalt not move thy plot location*** (Exodus 20:12(r)).

Biometricians, statisticians, and inventory designers should go the field at least once a year to keep reminded of what conditions are truly like. It is one thing to write procedures in the office and another to implement the instructions in the field.

One of the procedures that is particularly bothersome is the location of the sample plot when statistical sampling is used. The procedure writers make it emphatic that plots must be established at the location at which they were chosen. To locate the plot elsewhere, would bias the sample.

Once again I generally agree with this procedure, but there are at least two situations that I believe plots can be moved. The first is when the plot falls in a very uniform stand such as a plantation or lodgepole pine thicket. I have seen a great deal of time and money invested in surveying in the plot to be sure that it is exactly located at the place it should be. I have also seen a plot established by a contractor rejected because it was twenty feet off. Some type of logic must prevail in this type of situation.

The second case is when the plot occurs in an area that would endanger the cruiser. I have seen inventory crews risk life and limb to establish plots in some areas of the west. Again some kind of logic must prevail in this situation. No plot is worth serious injury or death.

A somewhat related problem is that of moving sub-plots when they straddle or cross a given condition such as an obvious stand boundary. Many of our Forest Service crews will move sub-plots of a ten-plot cluster, for example, if the subplots will fall out side the condition being sampled. In this case, I once again support the commandment providing there is no risk to life or limb. Edges of stands are part of the population, and not to sample them, ignores a portion of the natural world.

4. ***Thou shalt not exceed thy allowable sampling error*** (Exodus 20:16(b)).

Allowable sampling errors are frequently established by the inventory specialist or statistician to determine sampling intensity. Beyond this, the allowable and the resulting sampling errors appear to be seldom used.

I have been in the Washington Office of the USDA Forest Service for over long seven years. During that time, I have provided inventory statistics to the White House, Congress, news media, the Chief, and the general public. Never once during that time have I ever been asked about the reliability of the data provided. At the same time, I have obtained data from both Regions, FIA units, and other federal agencies and never once was I told what the associated sampling errors were (nor did I ask).

This is not to say that sampling errors are not important. It just points out that the only people who use them appear to be the inventory designers.

The Book of Revelation

And I saw an angel come down from heaven, having the key of the bottomless pit and a great chain in his hand (Revelation 20:1).

In the past few moments we have talked about platonic verses and inventory objectives. The establishment of inventory objectives is often relegated from God or the Chief Executive Officer to the Pharaohs or Staff Directors to the Data Collectors or Inventory Specialists. During this process, intents are lost and personal biases creep in.

Designs become customized to the unique perceptions of the individual that has the design responsibility. Those that establish the objectives, often have little or no experience in evaluating what is desirable or practical. In the end, the objectives may not be meaningful for the problems at hand.

In fairness to the writers of the scriptures, laws and direction are often purposefully kept general so that an organization has a wide latitude to maneuver. Without such latitude, there would be no need for people like me. In addition, if the laws were too explicit, and we didn't exactly follow the legislation, we would end up in court on a more frequent basis than we do now.

When you get a request for data or a request to launch an inventory program, use the questions listed above to clarify what it is that you are to do. Watch for hidden objectives. Look at designs and results of previous inventories for the same area to see if they are of any value. Lastly, examine existing data to see if they would suffice.

Remember that deliberate sampling does have a place in our inventories. Model-based and importance sampling will be discussed in depth during this meeting.

Don't be too harsh on crews that don't establish plots in exactly the right location. Examine the situation first before making them redo. Field crew leaders -- take your statistician to the field with you for a day or so, so that the biometrician can appreciate the real world situation.

As for sampling errors, there is an inconsistent axiom of which you need to be aware. The more urgent the need for data, the less concern there is about sampling errors; yet the more likely decisions based on the provided data will have a major impact on an organization's program.

There is another truth you should know. Allowable sampling errors are usually established by an inventory specialist usually high in the organization. If the inventory cannot meet the allowable error, consider "casting out" the specialist or changing the standards.

Finally -- ***Blessed are they that ask questions, consider the future, evaluate existing data, and plan accordingly: for they may have an inventory that provideth the sought-after answers and shall dwell at the right hand of the decision-makers forever.*** (Revelation 22:14(g)).

Let Me Give Thanks

I thank the following fellow apostles for their kind reviews and helpful suggestions for preparing these final scriptures: Jim Bones and Rich Birdsey, Forest Inventory and Analysis Branch; Milo Larson, Timber Sales; Karl Bergsvik, Timber Program and Management Planning Staff; and Dennis Murphy, Silviculture Staff. All are employed by the USDA Forest Service.

The Dead Sea Scrolls

Atterbury, Toby. 1987: Tutorial and notes for *Effective inventory: asking the right questions*. 1987 May 13-15; Corvallis, OR. Regional Workshop, Society of American Foresters. Beaverton, OR: Atterbury Consultants, Inc. Misc. Pagination.

Lund, H. Gyde. 1986a. A primer on integrating resource inventories. Gen. Tech. Rept. WO-49. Washington, DC: U.S. Department of Agriculture, Forest Service. 64 p.

Lund, H. Gyde. 1986b. Existing information and the sorcerer's apprentice. In: Oderwald, Richard G.; Burkhart, Harold E.; Burk, Thomas E.; eds. Use of auxiliary information in natural resource inventories. Proceedings: 1985 October 1-2; Blacksburg, VA. SAF 86-01. Blacksburg, VA: Virginia Polytechnic Institute and State University; School of Forestry and Wildlife Resources: 1-8.

USDA Forest Service. 1983. The principal laws relating to Forest Service activities. USDA Agriculture Handbook No. 453. Washington, DC: U.S. Department of Agriculture, Forest Service. 591 p.

Forest Inventory Purposes and Goals

Toby Atterbury - President
Atterbury Consultants, Inc.
3800 S.W. Cedar Hills Boulevard, Suite 225
Beaverton, Oregon 97005

ABSTRACT

Purpose and goals must be clearly defined to insure that a forest inventory will fully satisfy the needs of an organization. A written purpose defines the intent. Goals specify the timing, costs, uses, users, accuracy and other targets of the inventory. Each user must have input into the purposes and goals. Written purposes and goals are the most important part of forest inventory for they set standards for us to achieve.

INTRODUCTION

Forest inventory is one of the fundamental activities in the profession of forestry. Organizations can make forest inventory a part of the daily activities, with the information provided being used for appraisal, strategic planning, and operational planning. This creates happy, satisfied users. This can be achieved by using a process to insure success.

The most important and first step in the process of forest inventory is writing and agreeing on the purposes and goals. Written purposes and goals makes us into professionals and leaders by setting specific objectives to achieve.

DEFINITIONS

To help visualize these concepts Webster's New World Dictionary, Second College Edition defines the key words in this paper as:

Professional:
1. of, engaged in, or worthy of the high standards of a profession.

Leadership:
1. the position of guidance of a leader.
2. the ability to lead.
3. the leaders of a group.

Purpose:
1. something one intends to get or do; intention, aim.
2. resolution; determination.
3. the object for which something exists or is done; end in view.

Goal:
1. the line or place at which a race, trip, etc. is ended.
2. an object or end that one strives to attain; aim.

PROCESS

Successful forest inventories are accomplished by using a process which may be outlined as:

1. Define the purposes and goals.
2. Define the users and uses.
3. Define the products.
4. Choose the systems and methods.
5. Document the systems and methods.
6. Choose and train the people.
7. Collect the Data.
8. Process the data into useful information.
9. Analyze the information.
10. Report the information to the decision makers.

By far the most important step in this process is defining and writing the purposes and goals.

PURPOSE

The purpose of an inventory should be to produce results that can be used by the organization, specifically the users involved.

An example a forest inventory purpose statement may be: To provide (organization, client, owner, etc.) with forest inventory information in a way that timely accurate information is available for (appraisal, strategic planning, operational planning, etc.) so that better decisions can be made to enhance the organizations (profits, net present value, position).

Notice the purpose is only one sentence. It should be short and to the point.

The purpose statement must be agreed upon by all of the major users in the organization.

Purpose statements begin to create a vision of success for the individuals in an organization. In fact, this shared vision is the major ingredient for success.

GOALS

The goals must determine specific targets. They define specific dates, names, budgets, etc. Goals are measurable.

Example goals may be:

1. To have accurate inventory information (+ or -) of various items, such as volume, acres, etc.
2. To have the users believe the information is accurate and useful.
3. To finish the data collection and processing by December 31, 19xx, and have the information ready for the users.
4. To accomplish the inventory within the budget, $xxx,xxx.
5. To train the users to understand the information and use the computer programs.
6. To train the data collectors to a high degree of professionalism so that accurate information is collected.
7. To have the data and information available for use on miro-computers at the users desk.

8. To have the ability to add annual growth each year.
9. To maintain the maps and data for all changes due to management and natural causes.

Goals need to be specific, but should not contain procedures. Procedures should be done by the professional inventory forester. They should be changed as better equipment, knowledge, and information is available. Procedures should not become purpose, goals, or policy.

Procedures can become policy. Written purposes and goals should overcome this tendency and allow the inventory foresters to use various methods, techniques and procedures as needed. The methods, systems, procedures and people should be chosen that will accomplish the purposes and goals.

SUMMARY

Written purposes and goals make us professionals and leaders. They set standards for us to achieve. They give us drive, a reason to get up in the morning and work until late at night. They give us a way to measure our performance. They give us targets to aim at and hit. They give us reasons to come to these conferences, quest for more knowledge, improve our skills and inventory methods.

Written purposes and goals give us the ability to convince others we can accomplish something worthwhile. Accomplishment of the purposes and goals should bring us the most important reward, satisfaction of a job well done.

REFERENCES

Webster's New World Dictionary, Second College Edition, 1986.

OBJECTIVES OF A NATIONAL FOREST INVENTORY

FRENCH NFI MEETING USERS NEEDS

R.B. Chevrou
Statistician
National Forest Inventory Service
34970 LATTES - FRANCE

ABSTRACT

In the last years, French NFI service has made NFI data more accessible to all users through a huge database, queries and computer programs, easy to use from any place in the country.

This experience has shown what data, results and tools the NFI service should give to meet users needs. In short, they are :

Data, results and tools for estimating future probable crop, with general and detailed models for stands evolution ;

Wood qualities including stem form functions and computing tools ;

Various inexpensive volume tables from stem form functions ;

Height curves ; bark thickness ; with corresponding tools ;

Forest maps on computer files for location of data and results.

New theories are necessary for regression as most models and functions do not fit the usual assumptions for validity.

BACKGROUND

French NFI data are collected and results are published since the 60s. They are very interesting and useful to know the actual state of the forests, stand types areas, volumes and growths, per regions, species, age classes, and so on.

Little by little, a question has arisen : How to use these data and results for policy making at governmental level as well as at industrial one ? Many have tried to estimate forest resources, but they had neither an easy access to NFI data, nor efficient methods and tools for computational work, notably because, in France, almost half of the total crop comes from thinnings.

In the last years, we have built a system to meet users needs and two special tools have been developed which are of great importance : A database, including special computer programs in addition to usual queries, to make easy an access to NFI data from any place in the country through public network Transpac, and even in the world through existing networks ; and forest map files for locating NFI results.

So, it has been made possible to answer users questions, but not all of them yet, and NFI objectives can be changed to better meet users needs which are better known from the questions asked to the NFI database.

MEETING USERS NEEDS

The NFI database is located in a big computer Center (CNUSC in south of France) on an IBM 3090-400 mainframe, using DB2, PL1 and SAS. Its size is about 600 megabytes. For details, see GUERO M.C. and HOULLIER F. (1989).

From the actual state of the forests, easily obtained from usual queries, users try to meet their own goals.

Forest policy makers and industrials need estimates for forest resources in the next 20 years to come. It is of a great interest to get them in relation with stand types, species, qualities, accessibility, and location of the forests from a map.

Forest managers need rough data, volume tables, models for height and other variables, and stand types maps.

Forest Resources

Several different computer programs, included in the NFI database, give forest resources. In all cases, users have to interfere in the computational work by entering their own hypothesis for future management rules (thinning ratios, exploitable age or some other values related to it). These hypothesis are entered as answers to program questions.

Let notice that NFI data include data for past sylvicultural practices that help users choices.

From the NFI database, these computer programs give results with a classification for forest regions, stand types, species, diameter classes, qualities and accessibility.

Two computer programs are in use at the moment.

On a broad point of view : The user has to define an homogenous domain from NFI data ; It may be difficult to get simultaneously the homogeneity of the domain and a convenient ground plot number to get accurate estimates. Then the user gives 4 values ; The first of which, a mean diameter, is for classing stands or trees in two groups, the first group containing stands or trees with small diameters to be thinned, and the second for stands or trees with large diameters to be clear cut or regenerated ; The second value is a thinning ratio related to the above first group of stands or trees ; This thinning ratio includes natural mortality and it is expressed as a percentage of the annual yield ; For stands or trees to be clear cut, the user gives the maximum value of the annual cut and the minimum length of the period on which regeneration will last.

This method is called "Calcul des disponibilités forestières brutes" (DFB) that is "computation of crude forest resources". It is based on forest management methods used in France and it gives interesting results easily, but not the future state of the forests.

On an age class basis : the user has to define an homogenous domain for regular stands where ages have been measured ; then he gives, for each age class, the area to be clear cut and the thinning ratio for the rest of the age class area.

This method leads to entering many hypothesis (2 for each age class). With volume growth/ha in an age class seen as a constant through time, it leads to the state of the domain (areas and volumes in every age classes) at the end of a period of time equal to the age class width, so that computation can be done for another period of time.

Several other methods are in development and they already work on desk computer : A method dealing with diameter classes, either globally in irregular stands, or per age classes in regular stands in relation with thinning type ratio ; a method which modelizes NFI ground plot evolution, from which it is computed appropriate annual crop and the final state of the stands.

In relation with wood qualities, we have had to build special tools to estimate appropriate volume tables from NFI data as described below.

Overlaping an appropriate map file to the NFI map file permits to locate ground plots in watersheds or other territorial divisions, and then, from NFI database, to locate NFI results as well as resources estimates.

It has not often been done yet because most users do not see the power of such a tool to locate resources accurately and because the existing maps do not cover the whole country.

When the forest map is in a database (GIS), we can expect an easier access to this tool ; more-over, it will show which forests and stands types are of greater interest, so that the survey work may be increased for them and reduced outside.

Forest Management Purposes

When NFI started in France, good volume tables did not exist for all species all over the country. It has been decided to estimate volume on standing trees, measuring several diameters along the stem from bottom to top as shown in IFN (1985).

Bouchon, Delord and Rousseau (1986) have shown that measuring 4 diameters at convenient heights gives a very good tree volume estimate, with a relative bias less than 1% and a relative residual error less than 3%.

Delord (1984) showed that these 4 diameters give a good taper function modelizing tree stem form and leading to diameter estimate at any height with a good accuracy.

So, it is possible to estimate the volume of any part of the stem for given length and top diameter at any height.

Almost 2 000 000 trees have been so measured in the past on almost 200 000 ground plots and, since 1985, NFI uses its own volume tables from these trees. NFI results and forest resources can be computed with many different reference volumes using appropriate volume tables build from these trees.

A volume table computer program has been included in the NFI database, which can be modified on request to add new possibilities.

A similar program have been added to the NFI database for dbh bark thickness, which was measured on all trees and to day estimated using an appropriate function.

Forest managers and industrials can compute their own volume tables from these NFI trees when useful for a few dollars.

Each tree of the NFI ground plots is measured for radial growth in the last five years, and, when it is possible, height growth for the same period.

From these data, Chevrou (1986) has developed models for height and basal area, based on growth dx expressed as a function of the variable x (height, basal area) and of age a, with the following differential formula :

$$dx = f(x,a,c_i)$$

where c_i are parameters to be fitted to real data dx, x and a.

The computer programs work on desk computer and will be included later on in the NFI database.

Forest Map Files

In 1985, French NFI has changed its methods to get better area estimates by mapping stand types. It was seen that map files would be of very great interest for many purposes. They would make possible to use standard satellite data to make the forest map up to date every year or when necessary. It would be easy to add information from other files (with slopes, boundaries of watersheds or other territorial divisions, forest roads and trails). It would be possible to get from these map files, usual paper maps containing useful information and only useful one, that makes the map clearer and easier to read. It would make possible to modelize forest fires and to improve warning. Many other uses are not yet known.

Our main prospect is related to getting NFI results and forest resources for any territorial division related with governmental and industrial problems, or to road access. Using the forest map files gives ground plots numbers located inside the concerned area. From the NFI database, these plots lead to expected results. It will be necessary to build new routine estimates using outside plots as well as inside ones as it is done in mine prospecting. Using plots remeasured on several occasions, for instance SPR plots, should lead to very efficient estimates.

We plan to add other information to the forest map with density, height, or crown size values so that it will be possible to get volume estimates for each mapped polygon or each group of neighboring similar polygons, at least for the most important stand types. Our main purpose is to locate the stands to be thinned where crops will give mainly pulpwood, and to locate stands to be clear cut where crops will give mainly sawwood.

We plan also to put these map files in a huge database to make easier an access to these data and to link them to the existing ground plots NFI database.

Some Problems To Be Solved

Some problems have arisen in relation with the above methods used by NFI.

French NFI uses ground plots made by 3 concentric circles of radius 6 m for smaller trees, 9 m for middle size trees, and 15 m for bigger trees, as shown in IFN (1985). Such a plot is quite similar to point sampling plot with circle radius related to tree dbh. Lappi and Bailey (1987) have shown that such a plot lead to biased estimate for tree mean value on the plot and good estimates have to be found. Up to now, we work with mean values for dominant trees which often are located only on one out of the 3 circles, so that bias should be smaller.

Nevertheless, this NFI ground plot has some useful properties : It permits to measure appropriate proportions of small, middle size and big trees, excepted for very large trees ; boundary plots bring small bias as it is easy to use the mirror method or to enter the plot inside the proper stratum along a winding boundary.

Other problems are related to regression methods for fitting models and volume tables. Up to now these methods assume the model validity and we all known that we are working with invalid models. It would be of great interest to develop practical methods for invalid models and methods taking into account the correlation between volumes of trees located on same plots.

Working with forest map files leads to new concepts and new problems. We need easy routine formula to do what mine prospectors do in their own domain.

Conclusion

The NFI database, with its included computer programs, and the forest map files, permit users to work efficiently with NFI data. When these data and tools do not meet their requirements, new tools are built or new data are collected.

Such an interactive process between NFI and its users seems to be an efficient way to define the best objectives. Nevertheless, some special research work has to be done to find out future needs not yet clearly expressed, because some delay is needed before getting the useful ground data.

Though a very important work is done, specially in USA, on improving survey methods and on getting more accurate estimates, we feel that more efforts should be done to build routine methods and tools necessary to use efficiently survey data and results. Not only it might well be a good way to define better objectives, but also it could change our point of view about the data to be collected. It might well come out that objectives as well as data could be quite different according to the country and the prior knowledge that we may have in each case. Here, forest map would be essential, while there stem form would be of greater interest. In France, it seems that all aspects are of interest and it leads to a quite expensive national survey.

REFERENCES

Bouchon, J., Delord, J.M., et Rousseau, P., 1986, Comparaison de deux méthodes de cubage, RFF XXXVIII, 1, p 47-54.

Chevrou, R.B., 1986, Modélisation de l'évolution des peuplements à partir des données de l'IFN. 18ème Congrés mondial IUFRO. LJUBLJANA, Yougoslavie.

Delord, J.M., 1984, Le profil des tiges, RFF XXXVI, 6, p 505-514.

Guéro, M.C., and Houllier, F., 1989, La base de données de l'IFN, Nature et utilisation de ses données dynamiques, IUFRO workshop, S4-02-03/04, Forest Growth Data, Capture, Retrieval and Dissemination, GEMBLOUX, Belgique, April, 3-5, 1989.

I.F.N., 1985, But et méthodes de l'Inventaire Forestier National, Ministère de l'Agriculture, 65 p.

Lappi, J., and Bailey, R.L., 1987, Estimation of the Diameter Increment Function or Other Tree Relations Using Angle-Count Samples, Forest Science, vol 33, No3, p 725-739.

FOREST INFORMATION NEEDS OF CONCERNED GROUPS IN THE STATE OF MAINE

Don D. Tarbet
Forest Biometrician,
James W. Sewall Co
Old Town, Maine 04468
and
John H. Cashwell
Director,
Maine Forest Service
Augusta, Maine 04333

ABSTRACT

The State of Maine is the most heavily forested state in the US and is more dependent economically upon its forests than any other state. A wide variety of information is needed to make meaningful managerial decisions about this vital resource. Traditional inventories of standing timber and of growth do not completely fill the need, but must be considered as part of an overall structure of needed forest information. As culmination of an ongoing evaluation by the Maine Forest Service of forest information available in the State, representatives of the forest industry, landowners, government officials, and other interested groups were interviewed in 1988 to determine their needs for information on the forests of Maine. Differences in information needs appeared across user groups, but the general consensus was that information on growth and harvest were the most important components of a forest information system. Of these two, harvest or drain was regarded as the most difficult to study. From this study, conclusions were drawn concerning the types of information most needed in the management of the State's forests.

INTRODUCTION

Forest comprises about seventeen of Maine's eighteen and one half million acres of land. Ranging from predominantly spruce-fir in the north to a central hardwoods intrusion in the extreme south (SAF, 1955), this extensive forest cover is the dominant characteristic of the state.

The forest provides Maine and the nation with traditional values; jobs, recreation, air and water quality enhancement, and general spiritual benefits relating to quality of life. In economic terms, extractive uses of the forest add about four billion dollars to the gross state product. Recreation adds nearly another billion. (Maine Department of Conservation, 1988). This represents nearly 15% of the state's entire income, making the forest the largest single producer of revenue in the state.

The forest industry of the state is strongly oriented toward spruce and fir and has been affected by the recent attack of the spruce budworm. Also affecting the state's wood supply is a heavy demand for wood from the using industries, both in the state and in Canada.

The demand is so heavy that the US Forest Service survey of 1970 (Ferguson and Kingsley, 1972), projected that drain would exceed growth by the 1980s. Data from the USFS resurvey in 1980 (Powell and Dickson, 1984) suggests that the earlier prediction may be coming true, at least for some species.

At the very least, the margin where error can be tolerated in estimating supply and demand has almost reached zero. Gone are the days when growth and drain estimates could be very broad. Timber companies, lawmakers, woods workers, environmentalists, and everyone else concerned with the forests of Maine are aware of the problem. There is growing feeling that steps must be taken to protect this resource, but no hard evidence exists as to how serious the problem is now. No one knows whether the annual drain is less than the annual growth, or whether we have crossed the line between renewable resource utilization and mining.

Information on the state of the Maine forest comes from a variety of sources and is used in a variety of ways. Industry maintains some of its own data. The Maine Forest Service provides a series of reports on consumption, export, import, etc. The USFS resurveys provide estimates of standing timber and growth. These are supplemented by additional surveys.

The basic year-to-year drain figures are dependent upon reports from the wood-using industry and are reported by the Maine Forest Service. These figures are reported to the Maine forest service as mandated by law. However, no mechanism for enforcing the law exists, and those most familiar with the industry believe that the total cut is dramatically underestimated by these figures. The major reason for the underestimate is the under-reporting of exports. Various sources of information indicate that Maine's export figures are quite far below reality. Using the best information available, the authors have calculated that the total volume of wood cut from the forests of Maine may be underestimated by as much as 15-20%. In 1985 as much as a quarter of a billion board feet of timber may have left the state without appearing anywhere on drain figures.

The situation is clearly potentially destructive. The State of Maine needs a forest information system adequate to the job of not only defining the growth and drain on the forest but of modeling it into the future. The importance of adequate forest information systems has been considered by others (Baskerville and Moore, 1988). For Maine and for any other region for which the situation has become critical, two basic elements represent the ultimate expression of the state's information needs; growth and drain.

The inventory data of the US Forest Service are of questionable utility on any but the broadest scale also. As reported by Edson, Young, and Tryon (1985), errors are too high to make much use of the data in actual forest management and the time between resurveys is too long for potential users to have any confidence that the data represent actual conditions.

The question comes down to this: Can the forests of Maine continue to support the demands that will be put upon them, both for products and for quality of life enhancement? The responses are varied. The Ferguson and Kingsley (1972) projections sound a note of warning although the authors remain optimistic. Seymour et al (1985) listed potential timber shortfalls as did the Sewall Report (1983). Balter and Veltkamp(1987) also added warnings. Recently, however, Seymour and Lemin (1989) have struck a somewhat more optimistic note based upon modeling efforts.

It must be remembered, however, that all projections, are based upon data known to be flawed. An old computer expression is 'GIGO, Garbage In, Garbage Out'. The more sophisticated modeling schemes are most vulnerable to the inescapable fact represented by this old saw. The authors' experience indicates that a substantial part of the concerned community in Maine concurs with the above assessment of the available data and feels a need for information in which they can place confidence.

Drain, however, is dynamic. It changes with the economy, with perceptions of supply, with the weather, with the international market, etc. Figures collected every twelve years in a broad brush approach may not be adequate for dealing with the problem. This type of figure is the one most often considered sensitive by wood using industries, at least in the past.

Responding to the widely expressed dissatisfaction with current data on the forests, the Maine Forest Service contracted with the James W. Sewall Company in 1988 to analyze forest information needs in the State of Maine. As part of this effort, users and potential users of forest information were contacted to determine where the primary gaps and weaknesses in currently available information were perceived to be as a first step in developing a new forest information management scheme. This paper reports the portion of those results that relate most closely to measurement of growth and drain.

METHODS

In order to assess the quality of information now provided by the Maine Forest Service and to identify information needs of those concerned with the forests of Maine, potential information users were divided into the following groups: Forest Industry, Non-industrial landowners, Federal government, State and local government, service organizations (industrial support groups, consultants), and other interests (educators, private citizens, Maine Audubon, etc).

An attempt was made to obtain a representative list of individuals or organizations in these categories, but the impossibilities of defining actual sizes of groups or of randomizing the sample make it dangerous to generalize opinions beyond the limits of the sample. In all cases, the better known and more active representatives of the group were sought for interviews.

Seventy-one phone interviews were conducted with those representatives of the above groups considered key by the Maine Forest Service and by James W. Sewall personnel. In the interview, respondents were presented with a series of open-ended questions designed to elicit evaluations of the forest information now available and to obtain suggestions as to what the major information needs were.

Answers were then grouped in order to examine patterns of response. Most individuals gave more than one response.

RESULTS AND DISCUSSION

Table 1 lists the five most common answers given by the sample of respondents. It should be noted that numbers will not add up except across because each respondent was allowed multiple responses. No attempt was made to order the responses of any individual as to importance.

Most notably, four of the five most important perceived needs concerned, at least peripherally, the two basic elements described above as being major components of a forest information system, growth and drain. All 16 of those wishing for better information on imports and exports gave as their primary reason a need to know the magnitude of the drain. It is striking to note that over half of the industry representatives(10/19) focussed on this particular item.

The same is true in the case of the second most frequent response, that stronger legislation is needed in order to assure complete reporting of all data. Those who expressed this view referred to the need for obtaining drain figures. It should be noted that requests for this action were spread across all categories of information user.

More complete and more frequent inventory and better growth data were called for third (tied) most often. The inadequacies of the present inventory conducted by the US Forest Service have been discussed above. Here, again, most of the interest in inventory came from the forest industry. It is clear that industry is feeling concern about the state of the Maine forest and lacks confidence in the information currently available for making economic and managerial decisions regarding their own operations.

Table 1. Information needs by group.

		Group (number in group)					
Information needed	Total	Industrial	Owners	Federal Gov't	St Gov't	Service	Other
	(71)	(19)	(5)	(11)	(22)	(7)	(7)
Export/import	(16)	10	1		3	1	1
Legistlation	(13)	3		2	4	1	3
Inventory/growth	(9)	6	1		1	1	
Publish list of available data	(9)				7	2	
Mapping	(9)	1	1	1	4	1	1

Source: Sewall report (1988)

The next item (also tied for third), mapping of various elements of the forest included a call for land use mapping. Since land conversion to an unknown degree is occurring in Maine at this time, it is important to track these changes in order to understand the amount of growing stock in the state.

Clearly the two basic elements of a forest information system, growth and drain, are much on the minds of those concerned with managing and using the forests of the State of Maine. It is especially noteworthy the extent to which the forest industry representatives interviewed are concerned with these issues.

In the past, as stated above, it may have been possible for inventory and drain figures to be very loose. From what figures were available as recently as 1970 (Ferguson and Kingsley, 1972), growth greatly exceeded drain in the state. However since then, mills have expanded. Biomass has come into its own as a source of energy. The spruce budworm epidemic worsened and then tapered off.

It is reported that at least one sawmill is now accepting logs eight feet long and three and one half inches in diameter at the top end, wood that would have been considered inadequate for pulp 30 years ago (Young, 1989).

In spite of some contrary findings, it seems that there is at the very least a tightening of the wood supply in the state with its attendant affect on the entire economy as well as secondary effects on environmental quality.

In Maine where the forest has seemed such an extensive resource we are reaching the point reached earlier in other regions. The resource is being recognized as finite. Information to manage that resource is becoming more and more precious, especially to those most directly concerned.

Of the two basic elements, inventory and growth may be the easiest on which to achieve adequate information. The techniques of inventory are sufficiently advanced to allow a great improvement in current practices. The only difficulty is that inventory on the level that may be needed in the state is tremendously expensive. Establishing a large number of growth plots on top of other requirements may be prohibitively expensive.

For this reason it is becoming increasingly necessary to seek new inventory procedures, including those involving some form of remote sensing. Stripping ground inventories to the bone is another helpful approach. There are indications that under certain circumstances, height measurements add little to the accuracy of an inventory (Young, 1989).

Drain, however, is a difficult subject to approach. The growth rate changes slowly based upon known conditions. Inventory also changes slowly over time and can, if sufficient funds are available, be calculated. Drain is dynamic and changes due to fluctuations in the market place and to other factors that are difficult to identify, let along measure. While an interval of several years between measurements of inventory and growth may be acceptable, harvest levels are so subject to change that the intervals acceptable in the other case may not be acceptable here.

It is probably that remote sensing techniques will eventually handle questions of harvest, and it is hoped that efforts will be directed toward that end. Only with good data on drain can the equation become complete and a forest information system capable of meeting the increasingly critical needs of forest managers be achieved.

CONCLUSION

Most users of forest information in the State of Maine, especially those in the forest industry, feel that growth and drain data are severely lacking in adequacy for the task of managing the forest lands of Maine. Of the two, the drain figure is most glaringly weak. Efforts should be directed not only to improving inventory techniques, but toward perfecting techniques for estimating harvest amounts.

REFERENCES

Balter, K. and Veltkamp, J., 1987: Report on the demand for forest products in Maine. Prepared for: Maine Dept. Conservation, Forests for the Future Program. 70 pages.

Baskerville, G. and Moore, T., 1988: Forest information systems that really work, The Forestry Chronicle, April, 1988, p 136-140.

Edson, D. T., Young, H. E. and Tryon, T. C., 1985: A critique of statewide forest surveys, Unpublished manuscript, James W. Sewall Co. Old Town, ME, 9 pages.

Ferguson, R. H. and Kingsley, N. P., 1972: The timber resources of Maine, USDA For. Serv. Resource Bull. NE-26, 129 pages.

Maine Department of Conservation, 1988: Forest for the future, Report of Forests for the Future program, Maine Department of Conservation, Augusta, Me.

Powell,k D. S. and Dickson, D. R., 1984: Forest statistics for Maine -- 1971 and 1982, USDA For. Serv. Resource Bull. NE-81, 194 pages.

Sewall Co, J. W., 1983: Spruce-fir wood supply/demand analysis. Report prepared for: Maine Department of Conservation, 94 pages.

Seymour, R. S., Mott, D. G. , Kleinschmidt, S. M., Triandafillou, P. H., and Keane, R., 1985: Green Woods model: a forecasting tool for planning timber harvesting and protection of spruce-fir forests attacked by the spruce budworm, USDA For. Serv. Gen. Tech. Rep. NE-91, 38 pages.

Seymour, R. S. and Lemin, R. C., Jr., 1989: Timber supply projections for Maine, 1980-2080, Maine Agricultural Experiment Station Misc. Rep. 337, Orono, Me, 39 pages.

Society of American Foresters, 1955: Natural forest vegetation zones of New England, Committee on Silviculture, New England Section, Society of American Foresters, map.

Young, H. E., 1989, personal communication, James W. Sewall Co. Old Town, Me.

Using Forest Inventory Data in South Australia

J.W.Leech
Forestry Systems Officer
Woods & Forests Dept., PO Box 162, Mount Gambier, S.A. 5290, AUSTRALIA

ABSTRACT

Forest management practice in the state owned radiata pine plantation resource in South Australia is very intensive. The cut from the forest is set at the level the forest can sustain and there is little scope for expansion.

To meet the needs for sound planning, forest inventory has evolved considerably and has features not common elsewhere. The inventory procedure is described.

The inventory data are used to prepare short term (3-5 years) cutting plans which are the basis for scheduling harvesting operations. They are also used in the preparation of long term (commonly 70 years) predictions of future forest outturn and in the evaluation of alternative silvicultural strategies.

FORESTRY IN SOUTH AUSTRALIA

South Australia has been described as the driest state in the driest continent with only 1.2% of the land surface receiving on average more than 600mm (24") of rainfall a year. The native forests were generally poor and geographically restricted which led to the commencement of a plantation establishment program over 100 years ago.

Over 90% of the plantations are radiata pine and over 80% of these are in the south east of the state. This resource provides an allowable cut of about 700,000 m^3 (25 millon cu.ft. true volume) of timber per year to various government and privately owned utilisation plants including sawmills, a ply mill, pulp mills, particle board and preservation plants.

There is little scope for expansion of the plantation resource because of limited suitable land and competing alternative land uses. Estimates of present and future volume availability must be precise and unbiased if the allowable cut levels are to be set correctly.

PLANTATION SILVICULTURE

The historical development of radiata pine plantation silviculture in South Australia has been summarised by Boardman (1988) and the development of management planning techniques by Lewis, Keeves and Leech (1976).

Plantations are generally established with an initial stocking of 1600 /ha (650 /ac) using site specific weedicide and fertiliser regimes. Establishment is intensive with strenuous efforts being made to realize in practice the genetic gains from tree breeding programs.

At age 9.5 years all stands are assessed for site quality, a volume based index of site productivity, and this provides the primary variable for growth prediction and for defining thinning priorities. This serves as the first comprehensive inventory in a stand. We have found the use of volume based site quality superior to height based site index, especially for subsequent use in growth prediction models.

Stands are commercially thinned starting from ages 11-18 years depending on site quality and thinned every 6-8 years to residual stockings defined by Lewis (1963, Lewis, Keeves and Leech 1976). Stands are generally clear felled at between 45 and 50 years of age, the age being varied to balance wood flows by size assortments to the various utilisation plants.

Fertiliser may be applied during the rotation to increase productivity, generally not less than one year after commercial thinning if nitrogen is included in the fertiliser mix. Fertiliser may also be applied in anticipation of nutrient deficiencies.

After a 3-4 year fallow period stands are re-established.

Management planning is carried out using a computer based Yield Regulation system (Leech and Keeves 1981, Leech 1985).

Many Permanent Sample Plots have been remeasured for volume at 3-7 year intervals, and these measurements span over 50 years for a number of plots. They cover a wide range of thinning treatments and site qualities, and provide an unparalleled data base for model development.

INVENTORY

Inventory procedures have developed over many years but have changed little over the last 20 years. This has contributed greatly to the stability of the planning process.

Each plantation unit is divided into compartments of about 15-20 ha (37-50 ac) surrounded by an access track. Site variation and variations in past thinning history generally necessitate that each plantation unit is treated as a number of separate logging units. These logging units average 50-75 ha (125-200 ac), with the size decreasing with age. The logging unit can be considered uniform for the next commercial logging operation, be it thinning or clear felling. This unit obviously can vary over time as logging operations are carried out, but it is relatively stable and is the base unit for planning.

Inventory is by rectangular 0.1 ha (0.25 ac) plots established in each compartment in a stratified random manner based primarily on species, age and site quality, but also on standing stocking.

Inventory is compartment based rather than logging unit based as the unit cannot be defined with certainty until after the inventory has been carried out and volumes and stockings reviewed.

Plots are measured at least one year after each commercial thinning to reduce the error in the prediction of increment for stands immediately after thinning.

Plots are established at a rate of at least one plot per 4 or 6.5 ha (10 or 15 ac), with the greater intensity being used for areas where the next operation will be clear felling or later aged thinning, when it is important to predict the yield of large diameter sawlog more precisely.

Each plot is surveyed. Plot width depends on row width and plot length is adjusted to obtain an area of 0.1 ha (0.25 ac). The plot is surveyed in to compartment corners to facilitate re-establishment. Each tree is measured for diameter at breast height (1.3 m or 4'3"). Upper stand height, mean height of the 75 tallest trees per hectare (30 /ac), is measured on a small selection of plots where there is some doubt that the height can be estimated satisfactorily from height age relationships for the mapped site quality. The advantages of measuring more plots for diameter at breast height alone outweighs the advantages that measurement of upper stand height gives, particularly in regard to predicting yields by size assortments.

The inventory records trees that are expected to be removed in the next commercial thinning operation (thinnings elect). This is somewhat unusual in plantation inventory, but the extra effort has been found to be well worthwhile because variations in thinning type can be quite considerable between different areas. Thinnings elect trees in the plots are "marked" to the residual stocking level appropriate for the age and site quality at expected time of thinning, not at age of inventory (Lewis 1963, Lewis, Keeves and Leech 1976).

All data are keyed in the field directly on to hand-held Husky Hunter micro-computers (Leech, Sutton and Archer 1988) and the data down-loaded to a Vax 11/750 for processing. The use of the Husky Hunters has been readily accepted by the technical personnel as they are as fast or faster than manual recording, enable better manual and computer validation in the field, and hasten the updating of the master file.

There are some 7000 plots in the current inventory master file. These data are calculated and stored by the Inventory subsystem of the Yield Regulation system.

MODELS AND DATA

One of the strengths of the approach adopted has been the use of two different data sets in plantation management.

The Permanent Sample Plot series provides a pool of accurately measured data covering a wide range of thinning regimes (age of first thinning, thinning intensity, thinning type, and thinning interval), and of site quality, that can be used for model development. This may be augmented by other research plot data. The data base includes data that cover the complete rotation from before the age of site quality assessment until past the current nominal age of clear felling. Another important feature is that volumes have been estimated from sample tree volume measurements, either a 3 m (10') sectional method, or by a very satisfactory four-way tree volume table (Lewis, McIntyre and Leech 1973).

The inventory provides an estimate of the volume standing on each logging unit at present, and also the volume of the thinnings elect. Inventory plots are not remeasured to provide estimates of increment. Inventory plot data cannot be used to predict increment because they cover only one of many alternative silvicultural strategies. Increment is estimated using models based on the remeasured Permanent Sample Plots which cover a wide range of silvicultural alternatives.

YIELD REGULATION SYSTEM

The planning system or Yield Regulation system (Lewis, Keeves and Leech 1976, Leech and Keeves 1981, Leech 1985) has five main subsystems:

- Inventory: maintains a file of current measurements that provides estimates of standing growing stock and of the trees to be harvested in the next commercial thinning, including volume estimates by size assortments and estimates of short term increment.
- Area: maintains a file of current areas, according to their current usage and their present or future potential for wood production.
- Logging Category: enables logging units to be defined as groups of compartments or sub-compartments that are sufficiently uniform to be treated as homogeneous units in the future, and accumulates inventory data for these logging units.
- Cutting Plan: the Short Term Yield Regulation subsystem predicts volumes by log size assortments available from the next proposed operation and prepares cutting plans for the next 5 years. This is a tactical plan with only one silvicultural alternative generally being considered.
- RADGAYM II: the Long Term Yield Regulation subsystem (or RADiata Growth And Yield Model) predicts the volumes by log size assortments available from commercial operations in the future. The planning horizon is generally 70 years, covering at least a complete rotation and the transition into the next, but it can be extended to longer time frames for the evaluation of alternative silvicultural or economic strategies. Many silvicultural alternatives can be evaluated by this strategic planning model.

The system is fully implemented and operational. The first three subsystems which store and manipulate the data bases are quite stable.

SHORT TERM TACTICAL PLANNING

Cutting plans direct where logging is to take place for the next 5 years. They are generally reviewed and updated every 3 years, to facilitate contract logging, but the longer projection period allows better continuity of harvest scheduling.

Inventory plot data are used wherever possible. However for the first commercial thinning, estimates are generally based on site quality, standing stocking, expected stocking after thinning, extraction row frequency and age. This has been found to be satisfactory as the operation produces small wood used for pulp wood and preservation material.

The system allows for increment between time of inventory and time of thinning to be modelled, especially for the thinnings elect. This increment can be quite considerable and the volume to be removed has been found to double in a five year period between inventory and subsequent thinning for high site quality stands.

Volume estimates are produced by size assortments including volumes to 7, 15, 20, 25, 30, 35, 40, 45 and 50 cm top diameter underbark limits (ie 3" and 2" log diameter classes from 6" to 20").

Volumes are estimated in cubic metres, true volume, underbark. Only one silvicultural regime is generally considered in this planning phase.

Most stands have been thinned for many years to the Optimum Thinning Range (Lewis 1963, Lewis, Keeves and Leech 1976) and the age class distribution is broadly uniform, so it is generally possible to schedule second and later thinning operations at the time they are silviculturally due. Clear fellings are used to balance the wood flows by size assortments in order to meet plan objectives. First thinnings are manipulated as far as possible to meet the requirement for small diameter pulp wood and preservation material. Limited areas of pre- and non-commercial thinning have been carried out to prevent a serious backlag of first thinning from building up, and thereby improving stability to wind and increasing tree size.

The system is stable and has only undergone minor modification over the last 10 years.

LONG TERM PLANNING

The Long Term Planning Model or RADGAYM II (Leech 1985) is continually being modified to test various silvicultural options for the future.

The system produces similar outputs to the Cutting Plan subsystem, but results are commonly summarised as five year averages and presented graphically. Given the same input the two subsystems are consistent.

The planning horizon usually evaluated is 70 years, covering the current rotation for all forest either planted or planned to be planted, and also the transition into the next rotation. This was especially necessary to incorporate the replanting backlag after the devastating forest fires of 1983 (Keeves and Douglas 1983). For some analyses the horizon can be extended much further.

The subsystem simulates the growth of each individual logging unit over the planning period, and accumulates the results. Future operations are scheduled according to predetermined rules but these can be overridden for a particular logging unit.

The subsystem produces a summary of the volumes by log size assortments that are expected and also produces cash flows using a table of current costs and royalties and calculates net present value at a variety of real interest rates.

There are two quite different uses for the subsystem. Firstly, present and proposed practices can be evaluated. A combination of stand and forest based analyses in the past indicated the desirable silvicultural strategy and this has been accepted as the base model or proposed practice model. This implies a target forest, and the transition from the current state to a near steady state some time in the future. The objective is to supply wood to various utilisation plants, as required by log size assortments and quality categories; and that implies profitability, near maximum productivity and non-declining wood flows.

There is limited opportunity to markedly change the silvicultural practices used because the allowable cut is set at the maximum level that the forest can sustain and is completely committed to the existing industry. There is little likelihood of expanding the area of plantations. Any marked change in silviculture would have to be implemented carefully over a considerable period of time. There is no legislative requirement for non-declining wood flows, just the implicit recognition that the Government owned forests must be managed to enhance industrial stability and for the benefit of the people of South Australia.

Growing stock will vary considerably during the long planning period as volume production by size assortments is balanced to provide smooth levels of wood flows. The variation from the average can be up to 10% for total standing volume, 20% for sawlog, and up to 30% for large diameter sawlog.

The second use of the system is to investigate alternative silvicultural or management strategies, in order to increase productivity. Leech (1988) presented some 40 alternatives to the proposed practice model. These are a subset of over 200 alternatives evaluated and demonstrate the effect of changing thinning regimes (age of first thinning, thinning interval, thinning type, and thinning intensity) and of changing the objective rotation length, as well as other alternatives. This forest based analysis implicitly takes into account the transition from current practice to the alternative practice. These alternatives are analysed in terms of net present worth and in terms of wood flows by size assortments.

Over the years the proposed practice model has changed only gradually and minimally, and the approach adopted gives forest based control in the planning process. It also allows decisions about options such as changing rotation length, or fertiliser policy, to be based on analyses of the forest and not simply of single stands.

It is interesting to note that some alternatives considered desirable when based on stand analyses were found to be unsatisfactory when the forest based analyses were considered. One example was that stand based analyses show that the optimum time to apply fertiliser is after the second thinning but forest based analyses indicated that it was best to apply fertiliser after every thinning for certain age classes and not apply fertiliser to other age classes at all.

It has been decided not to expand the system to include a mathematical optimisation phase. Instead the model is run with a wide range of silvicultural and management options and the response surface examined to determine the desirability of changing these options. This allows a number of objectives to be considered. In our view this gives much better control of the planning process.

It also enables alternatives that may have unquantifiable features to be arbitrarily considered in the decision making process.

The system is fully operational but changing the alternatives to be evaluated may require changes to the program. It would appear desirable to have a fully interactive system that could cope with all possible alternatives, but this has been considered to be too expensive at this stage.

Great care has been taken in implementing this subsystem, with the objective being to produce a simple, modular and flexible system.

SUMMARY

The Woods and Forests Department in South Australia has developed a soundly based system that enables it to effectively regulate and control the intensive plantation enterprise now being managed as a mature resource.

The system meets the three most important criteria; it is complete, it is operational, and its reliability is accepted.

REFERENCES

Boardman, R. 1988: Living on the edge - the development of silviculture in South Australian pine plantations. Aust. For. vol.51, no.3, pp135-156.

Keeves, A. and Douglas, D.R. 1983: Forest Fires in South Australia on 16 February 1983 and Consequent Future Forest Management Aims. Aust. For. vol.46, no.3, pp148-162.

Leech, J.W. 1985: A management planning model of a large plantation forest. Mathematics and Computers in Simulation vol.27, pp199-206.

Leech, J.W. 1988: Analyses using the South Australian Long Term Planning Model. Proc. Workshop "Modelling Trees, Stands and Forests", Melbourne Victoria, August 1985, Bulletin 5 University of Melbourne, pp445-459.

Leech, J.W. and Keeves, A. 1981: Strategic and tactical planning in South Australia. Proc. I.U.F.R.O. Project Group P4:02.1, "Economics and Techniques of Thinning Plantations in Australia and New Zealand." Canberra A.C.T., 4pp.

Leech, J.W., Sutton, M.W. and Archer, G.R. 1988: Implementing field data entry programs on Husky Hunter micro-computers. Proc. Research Working Group No. 2: Mensuration and Management, Gympie Qld., September 1988, 9pp.

Lewis, N.B. 1963: Optimum Thinning Range of Pinus radiata in South Australia. Aust. For. vol.27, no.2, pp113-120.

Lewis, N.B., Keeves, A. and Leech, J.W. 1976: Yield regulation in South Australian Pinus radiata plantations. Woods and Forests Bull. 23, 174pp.

Lewis, N.B., McIntyre, G.A. and Leech, J.W. 1973: Regional volume table for Pinus radiata in South Australia. Woods and Forests Department Bull. 20, 135pp. (Metric edition of Lewis, N.B. and McIntyre, G.A. 1963: Woods and Forests Department Bull. 14).

Contracting for Inventory Plot Measurement: the Minnesota Experience

D. Heinzen and S. Flackey
Resource Assessment Unit
Forestry Division, Minnesota Dept. of Natural Resources
2002 Airport Road, Grand Rapids MN 55744

ABSTRACT

Between the 1977 and the current Forest Inventory and Analysis cycles in Minnesota, the Department of Natural Resources turned from use of state personnel to employment of outside forestry contractors for precision measurement of field plots. Results have been highly favorable in terms of worker productivity, survey costs, and data reliability, but the change has required new initiatives and great flexibility on the part of the agency with regard to contractor development, training, plot acceptance sampling and contract administration. The lessons learned have significant implications in an era of straitened agency budgets and increased reliance on contracting.

INTRODUCTION

The Minnesota Department of Natural Resources established our Resource Assessment Unit (then called Forest Inventory) in 1974. From then until 1986, the unit used permanent and temporary state employees to perform the permanent plot measurement tasks entailed by the state's participation in the Forest Inventory and Analysis (FIA) program coordinated by the North Central Forest Experiment Station. Since the mid-1980s, worsening budgetary constraints have dictated that our unit find ever more cost-effective means of getting fieldwork done.

An alternative that had been considered at the outset was contracting for plot measurements. The idea had been greeted with considerable skepticism in various quarters. Contractors, it was felt, would not produce the consistent, high-quality data that could be expected from state employees. Despite the prevalence of this attitude, the budgetary situation persuaded us that contracting was an option to be explored.

Through contacts with other agencies we obtained bidder lists and sample contracts, and used these to develop our own. We awarded our first contract in June 1986, and since that time have carried out most of our fieldwork in this manner. However, we continued to conduct some fieldwork with permanent and temporary employees, and thus we have a basis for making direct comparisons. In what follows, we describe the methods used and evaluate the alternatives in terms of cost and data quality.

METHODS

Permanent employees were based at five separate locations to avoid per diem costs. Crews consisted of two persons, at least one of them trained and experienced in inventory work. Four-wheel-drive trucks, all-terrain vehicles and snowmobiles were supplied. Wages were based on an hourly rate, with actual work schedules left to each crew's discretion.

Temporary employees were based at one location and received per diem allowances in travel status. One-person crews were used, but crewmembers could choose to work in pairs if the nature of the plots warranted it. Crew supply and equipment was on the same scale as permanent employees'. Two of the temporaries hired were already experienced in inventory; the others attended a one-week training session.

Contractors submitted sealed bids for packets of plots. Eligibility for bidding required six months' inventory experience or a 2-year forestry degree. At least one of the crewmembers measuring any given plot had to be certified either by passing a test or by successfully completing our one-week training session. Contracts ran for eight months on the average, with three-month extensions granted if necessary. Contractors submitted their work in blocks of 10 plots. One plot in a block was randomly selected for acceptance testing; it was visited and completely remeasured by a Resource Assessment forester. If the check plot achieved a passing score of 75 (out of 100 points), payment for the block was made. If it failed, the remaining nine plots were returned for re-examination by the contractor--or, if circumstances warranted (as in case of fraud) we could terminate the contract and take the performance bond.

COSTS AND ACCOMPLISHMENTS

Table 1 summarizes the accomplishments of permanent, temporary and contract crews in terms of cost per plot. Permanent crews were used for 15 months between June 1986 and August 1987 and averaged 76 plots per person-year at a cost of $370.05 per plot. If per diem costs had been incurred, cost per plot would have risen to about $450 for permanent employee crews.

Temporary employees were utilized for nine months between October 1988 and July 1989, and also averaged 76 plots per person-year. Their cost per plot came to $341.46. The higher cost of permanent crews arose from salary and fringe benefits, partially offset by the fact that temporary employees averaged 65 per cent of their time in travel status earning per diem. During the nine months two of the temporary employees were terminated on account of poor production.

From July 1986 to the present, contracts for 7043 plots were awarded to 30 different contractors, at an average bid of $155.48 per plot. Some 4047 of these plots have been completed and constitute a basis for calculating an additional $89.68 per plot in administrative costs, including the cost of acceptance checks. This brings the average total cost per contractor plot to $245.16.

Table 1. Cost of FIA inventory plots.

Crews	Time Period	Plots	Outlays	Cost Per Plot
Permanent	June 86-August 87	435	1268 days @ $11.02/hr plus fringe & equip. =$160,973.11	$370.05
Temporary	October 87-July 89	184	$47,123 salary 7,168 per diem 8,500 equip. =$62,811	$341.36
Contract	July 86-July 89	7,043	$1,095,087 bids	$155.48 bid 89.68 admin $245.16 total

DATA QUALITY

The assumptions on which the U.S. Forest Service bases the design of FIA inventory require that the complex set of navigational and mensurational tasks involved in taking each 10-point permanent plot be carried out with repeatable precision. Deviation from a high standard of accuracy can seriously compromise the inventory, and the effects of poor execution can persist for decades. Regardless of who performs inventory fieldwork, some form of inspection is clearly necessary. A significant advantage of the contracting approach is that quality-checking is built into the payment mechanism. We are forced by the contracting system to carry out inspections in timely fashion; moreover, our checks must be thorough, quantitative and (if need be) defensible in a court of law. Where state employees crews are used and no one's income is immediately at stake, other priorities can easily interfere with quality control. So we found it in our own efforts to compare results. We had intended to inspect permanent, temporary and contract crews on an identical basis; but "other things came up" and we ended by inspecting only 2.8 per cent of the permanent crew plots and 3.3 per cent of the temporary crew plots. But we inspected a full 10.4 per cent of the contract plots--we had to, the contractors' pay depended on it.

On our acceptance checklist the inventory crew starts with 100 points and loses points for substandard performance in any area of plot location and measurement. In FIA design, some plots require more complex measurements than others, but regardless of complexity no more than 25 points may be lost. Excluding all non-passing scores, permanent employees averaged 86.9 points, temporary employees averaged 86.0 points and contractors averaged 85.3 points. However, including failed inspections in the tally changes the result substantially: permanent employees then average 62.0 points, temporaries 75.7 points, and contractors 82.9 points. Permanent crews failed 75 per cent of their inspections, temporaries 33.3 per cent, and contract crews only 5.7 per cent. This gross disparity probably arose from the fact that contract crews who failed inspections had their work returned for re-examination, while the same did not hold true for state employees. It is likely that the quality of state employee output would have improved if failed inspections had resulted in returned work, but it is also likely that costs per plot would have risen.

ADMINISTRATION

"Administration" is a vague term, and its content as applied to the costs of running an inventory program is matter for argument. The per-plot cost figures given above include a significant administrative element for contract crews only. Table 2 shows the components of this calculated administrative cost. The total outlay incurred to date is divided among the 4047 completed plots to arrive at the per-plot cost of $89.68. This figure is somewhat inflated, as some of the incurred costs pertain to the nearly 3000 plots not yet completed. More important than this is the fact that significant administrative outlays should in fairness be added to the cost of plots measured by state employee crews as well. Of course none of the costs peculiar to the contract-issuing process would apply; but approximately $30 of the total administrative cost per plot arises from acceptance checking, and this (as our data show) is at least as important with state crews as with contract crews. Keypunching and editing cost about $15 per plot regardless of who has gathered the data. In addition, state crews must receive inventory training, and also require additional overhead expenses in hiring, supervision, time reporting, performance reviews and the like, which together probably amount at a conservative estimate to $5 per plot. Thus approximately $50 in administration should be added to state crew costs to place them on a basis of equality with contract costs. This leaves us with a permanent employee cost of $420.05 per plot, a temporary employee cost of $391.36 per plot, and a contract crew cost of $245.16.

Table 2. Administrative items and costs.

Tasks	Outlays
Write/revise contracts	
Obtain/maintain bidder list	
Mail bid invitations	
Respond to inquiries on prospectus	
Copy contracts	
Mail contracts	
Bid openings	
Rebids	
Contact bidders for acceptance	
Process contracts	
Training	
Complimentary inspection	
Acceptance inspections	
Process payments	
Contractor correspondence, inquiries	
Reprocess seized & forfeited contracts	
TOTAL, Labor & Supplies:	$236,024
Fiscal year lost opportunity cost	
Keypunching	
Error checking & editing	
TOTAL, Labor & Supplies:	$126,906
	$362,930

$362,930 / 4047 completed plots = $89.68 admin. costs per plot

RESULTS AND DISCUSSION

Our data thus seem to show without much ambiguity that contract inventory crews gather inventory data at least as well as state crews at lower cost per plot. While this is our chief conclusion, some qualifying observations should be made. Managing an inventory program on a contract basis opens an agency to some unaccustomed frustrations and delays in addition to the opportunities. Not all contractors are successful. Some 17 per cent of our contractors were terminated by us for various reasons, and an additional 13 per cent forfeited their contracts. These represented about 16 per cent of all contracts issued, and more such problems may occur before the present inventory cycle is complete. If this sort of thing is unexpected and unplanned for, the attrition can cause serious delays in data availability. Its impact can be minimized by the exercise of forethought and by close contact with contractors. Inventory unit staff must be available to contractors in off hours; they must learn to be flexible yet consistent in their dealings with contractors. The contract itself must be written to emphasize quality and timeliness, without imposing unnecessarily restrictive conditions (such as numbers of personnel per crew) on how the contractor accomplishes the job. Before rushing into contracting, an agency must reorient its thinking towards questions of vendor development, contracting cycles, and the absolute need for competent acceptance checking on a timely basis. Failure to address these matters will lead to serious trouble.

We cannot conclude without wondering out loud why our generally successful experience with contracting inventory measurements ran counter to all the expectations--or rather, why the expectations were so pessimistic in the first place. Did the generally low prestige of inventory work have something to do with the skepticism, or was it just garden-variety institutional conservatism?

Section 2

Forest Inventory and Sampling Designs

Moderator: Timothy G. Gregoire

William G. Warren, Line Intersect Sampling: An Historical Perspective

W.E. Frayer, Inventories of U.S. Wetlands

Thomas W. Beers, Precise Tree Frequency Estimates in Horizontal and Vertical Point Sampling

Walter Bitterlich, A Comprehensive Concept for Future Forest Inventories

H. Gary Murchison and Alan R. Ek, Efficiency of Two-Phase and Two-Stage Sampling for Tree Heights in Forest Inventory

Stephen A.Y. Omule and David W. Ormerod, Sampling for Logging Residue and Waste in British Columbia

Kim Iles, Critical Height Sampling: A Workshop on the Current State of the Technique

Victor A. Rudis, Sampling and Modelling Visual Component Dynamics of Forested Areas

John C. Rennie, Model-Based Sampling

Harry T. Valentine, Timothy G. Gregoire, and George M. Furnival, Importance Sampling for Volume With a Portable Computer

Line Intersect Sampling: An Historical Perspective

William G. Warren
Research Scientist
Science Branch, Department of Fisheries and Oceans
P.O.Box 5667, St. John's, NF, Canada A1C 5X1

ABSTRACT

The development of line intersect sampling in forestry is reviewed. Topics include the motivation for the method as originally conceived for the assessment of logging residue, the modifications that have been suggested and employed by various workers, the evaluation of performance including the determination of the variance of estimates so obtained, the handling of non-random orientation of pieces and the relationship of the method to other sampling techniques, in particular, angle-count (horizontal-point) sampling. The paper concludes with some general thoughts on the place and role of line intersect sampling.

The Origin of Line Intersect Sampling

Line intersect sampling has its roots in Buffon's (1777) needle problem. Although Canfield's (1941) extension of line transect sampling and Matérn's (1964) method for estimating the total length of roads by means of a line survey contain something of the essence of line intersect sampling, the method as currently applied in forestry stems, primarily, from Warren and Olsen's (1964) development for the assessment of logging residue. As far as is known, it was these authors who first used the term "line intersect sampling". Similar approaches have been used in other areas, for example the estimation of fiber length in microscopically observed pulp samples.

In 1958 in New Zealand, Warren and Olsen were faced with the problem of determining an economic means of estimating the volume of logging residue following clearcutting. The terms of the log sale provided for royalty payments on all merchantable material, where merchantable was defined as a minimum of four feet in length and four inches small end diameter. The logging company believed it to be more economic for them to extract only pieces that exceeded a somewhat larger limit, and to pay the royalty on the remainder. The task was then to determine how much of the smaller material was being left on the site.

The local forester naturally started with conventional (1/4 ac.) circular plots, but found that the interplot variation was so great that a large number was required to obtain estimates with acceptable precision. Further, the searching of the plots was time consuming and, unlike standing trees, many pieces traversed the plot boundaries. Some improvement was achieved by the use of long narrow plots although, in relation to the value of the material, these were still found to be uneconomic.

However, the reduced variation obtained with the long narrow plots, coupled with the notion of exploiting the theretofore inconvenience of boundary intersections, led to the idea that there should be a relationship between the number of pieces intesected by a line run at random through an area and the volume contained within that area.

Warren and Olsen derived the formula $V = 660\tilde{v}/I_cL$ where V is the volume measured in cubic feet per acre, L the length of the sampling line in chains (1 chain = 66 feet) and $\tilde{v} = \Sigma\, v_i/l_i$, where v_i and l_i are, respectively, the volume (in cubic feet) and length (in feet) of the i^{th} piece intersected by the sampling line. I_c is a factor that depends on the orientation pattern of the residue. If the pieces in the area are randomly oriented, i.e. the distribution of the angles that the pieces make with the sampling

CLEMSON UNIVERSITY LIBRARY
PUBLIC DOCUMENTS DEPT.

line is uniform on $(0,\pi)$ then $I_c = 2/\pi$. In tractor-hauled areas this seems a viable assumption. Warren and Olsen also used an empirically derived 0.67. A larger value is required for cable-yarded or steep-slope operations if the sampling line is at right angles to the predominant hauling direction or slope. Such situations will be discussed further below.

Warren and Olsen also showed that $\tilde{v}$ and n, the number of pieces intersected, are related, in expectation, by a factor that depends on the piece-size distribution. They also found that, because they were dealing with basically single-species, even-aged plantations logged by the one company to the same utilization standard, this factor was notably stable about a value of approximately 0.33 ($= \alpha$). Accordingly, their final formula took the form $V = 660\alpha n/I_cL$. Since only a count of the number of intersected pieces was required to give estimates comparable to those obtained from parallel plot sampling, the procedure provided the needed economic tool. Although the method was put into practice in 1958, publication was delayed until 1964 because of Dr. Warren's absence in the United States and Canada from early 1959 to late 1962.

Relationship to Other Methods

It should be clear that a count of the number of pieces intersected yields an estimate of the total length of pieces per unit area; one need only replace $\tilde{v}_i = v_i/l_i$ by $l_i/l_i = 1$. It is no accident that this is reminiscent of Bitterlich sampling for basal area per acre. Line intersect sampling (LIS) and angle count sampling are just two manifestations of the same general principle; they are variations of Poisson sampling (*sensu* Hájek, 1964) as also, incidentally, is Grosenbaugh's (1964) 3P sampling. This is made clear in the elegant theoretical development given by De Vries (1973, 1979a, 1986). One can, however, build a bridge between LIS and angle-count sampling by considering, firstly, instead of logs in a clear-cut area, standing trees and imagining these felled at right angles to the sampling line. One would then have LIS with $I_c = 1$. This has been termed vertical line sampling and should be viable in sparsely stocked stands. It is then a short step to Strand's (1958) horizontal line sampling and, thence, to horizontal point (i.e. Bitterlich or angle count) sampling.

Modifications

The next contribution came from Van Wagner (1968) who, because of his involvement with forest fire research, had the objective of estimating the quantity of slash or larger-sized fuel on the ground. He provided an alternative derivation which, depending on one's point of view, might legitimately be described as either a generalization or a particularization of Warren and Olsen's development. Note that $\tilde{v}_i = v_i/l_i$ can be considered an average cross sectional area of the i^{th} log. If all the logs were cylinders, as assumed by Van Wagner, $\tilde{v}_i = \pi d_i^2/4$ and, accordingly, the expression for volume per acre, under random orientation, becomes, apart from a constant depending on the units of measurement, $V = \pi^2 \Sigma d_i^2/8L$.

Van Wagner, of course, recognized that logs are, in general, tapered but argued that the estimate would still be unbiased. He addressed the orientation problem by showing that the maximum possible error (which occurs when all the pieces are aligned in the same direction) was substantially reduced by running a line at right angles to the first, and was close to being negligible with two additional lines at 60^o to the first.

Warren and Olsen's approach to the orientation problem was also to use lines at right angles. Let I_c be the orientation factor appropriate to one line and I_s the factor appropriate to the other. They showed that, if n_c and n_s were the numbers of pieces intersected, I_c/I_s could be estimated as n_c/n_s (appropriately adjusted if the lines were of different length) and that the relationship between I_c/I_s and I_c was surprisingly stable over the choice of a form for the distribution of the orientation angle. Thus, from n_c and n_s one could estimate I_c and I_s, and the two estimates so obtained could be averaged to form the estimate for the area. A requirement is that the angular orientation of the pieces is independent of their length. Warren and Olsen's field studies confirmed the viability of this assumption.

There is little information as to which approach is the better. Pickford and Hazard (1978) cite Bailey (1969, 1970) as finding that, for populations with nonrandom orientation, the assumption of uniform distribution was justified, provided two sample lines were run at right angles to each other. However, Bailey (1970) shows a clustering of his I_c/I_s v. I_c data towards 1.0:0.637, implying that, in his test areas, the tendency was much more towards random orientation than towards alignment. As far as is known, no comparisons have been made under a tendency towards alignment.

Bailey (1970) also compared the "intersected" area, i.e. the cross sectional area at the point of intersection (after Van Wagner, 1968), the mid-length cross sectional area (on 4 of his 13 sites) and the mean of the end cross sectional areas; the last mentioned would give the correct piece volume under the the Smalian formula, on which the B.C. Cubic Scale was based. He found little difference between the use of mid-section and intersected areas but concluded that these were biased relative to the mean of the end areas. However, Pickford and Hazard's (1978) simulation study apparently confirms Van Wagner's contention that taper does not introduce bias. The discrepancy lies in the choice of the Smalian formula; for the frustum of a cone, the volume is overestimated by $(\pi\ 1/4)(d_2 - d_1)^2/6$. The problems of volume definition are well described in De Vries (1979a, 1986). Clearly the use of the intersected area reduces measurement time, even in relation to the mid section area (Van Wagner and Wilson, 1976) but presumably at the expense of increased variability in the estimate (confirmed by Bailey, 1970, and Pickford and Hazard, 1978).

Precision

De Vries (1979a, 1986) derived the theoretical variance of the estimate. In this respect there is a subtle difference between Buffon's needle problem and LIS. In the former, needles are independently cast at random onto a lined area. Thus the probability of an intersection for any one needle is independent of the probability of an intersection of any other needle. In LIS, however, it is the needles that are fixed and a line placed at random. Knowledge that a certain needle has been intersected then affects the probability of intersection of the remainder. This introduces a covariance term that cannot, in general, be specified exactly although, given the scale of logging residue applications, this component is almost certainly negligible. A second approximation is introduced by taking $p(1-p) = p$ since p is small. It often appears to be overlooked that Warren and Olsen derived an expression for the coefficient of variation of V which, when transformed, is equivalent to the expression for the variance given by De Vries. These expressions for the variance involve, of course, the unknown population values. In practice the precision is estimated through the variance of the volume estimates generated by more than one sampling line.

Pickford and Hazard (1978) reported some fairly extensive Monte Carlo simulations of LIS to determine its precision under various population densities, and size distributions including taper. Many of their results could have been obtained by application of the formula given by Warren and Olsen. Although there is no overlap, the results of the situations considered by Pickford and Hazard are consistent with those presented in Table 1 of Warren and Olsen.

In a follow-up paper, Hazard and Pickford (1986) explored, again by Monte Carlo simulation, the effects of different sampling strategies on tractor and cable yarded populations, the former appearing randomly oriented and the latter with a strong tendency to alignment. They used a single sampling line, two lines at right angles and three lines at 60^o, both systematically and randomly oriented. Interestingly, for the cable operation, they found the bias with the two lines at right angles to be somewhat less than with the three lines. Although their preferred system turned out to be two lines at right angles randomly oriented, overall the variations appeared as remarkably robust, unless something unintelligent was done, such as using a single sample line roughly parallel to the direction of piece alignment. They also expressed the view that Warren and Olsen's approach, particularly if the area were also stratified, would likely yield estimates as precise as their preferred system, but the method was not incorporated in their study.

Applications

Various other expositions on the subject have appeared since Warren and Olsen's paper in 1964. Most were directed towards evaluating the operational feasibility of the method in particular circumstances, with modifications sometimes incorporated. Of the more notable is the "planar intersect" method of Brown (1971) and Brown and Roussopoulos (1974). Other works include Howard and Ward (1972) with an early attempt at evaluating alternative sampling strategies, Ménard and Dionne (1972), Martin (1976), and Meeuwig et al. (1978). This list is by no means comprehensive.

De Vries (1974) detailed the theory of multi-stage line intersect sampling. Apart from the example given therein, based on ficticious data for estimating the number of bark beetles per acre, as far as is known no applications have been made.

Further Thoughts

Line intersect sampling is essentially a means of selecting a sample and deriving inferences concerning a population without setting up a frame in the conventional sense. As Bitterlich sampling provides a natural measure of stand basal area, LIS provides natural estimate of the total piece length per unit area. Likewise, as angle-count sampling requires the measurement of the basal area of all selected trees for the estimation of any other property, LIS requires the measurement of the length of the intersected pieces. The exception is volume which may be estimated from the measurement of the cross sectional area at the point of intersection. This can be done without the measurer moving from the sampling line. In contrast, with Bitterlich sampling, the observer has, in general, to move from the sweep center to every selected tree. Indeed, customary practice requires the measurement of the distance to, and the diameter of, borderline trees, although the merit of so doing may be questioned. Provided the observer is not biased, the misclassification of borderline trees simply adds to the variance of the estimate, and may be more than compensated by a few extra sweep centers at less cost than the routine checking of borderline trees. Indeed, when it comes to something other than the natural measure, i.e. basal area, there is no reason to suppose that this means of sample selection would have any advantage over other methods, including conventional plots. Of course the same argument applies to LIS although, perhaps, to a lesser extent, since the log on which the measurements are to be made is, so to speak, at the observer's fingertips. Nevertheless, this may explain why multi-stage line intersect sampling has received so little attention.

Before closing it seems pertinent to emphasize the difference between line intersect sampling, line intercept sampling and line transect sampling. The distinctions are well drawn by Pielou (1985). Line intersect sampling, which Pielou also calls Buffon sampling, is essentially a method of unequal probability sampling that does not require a frame in the conventional sense; any attribute of the sampled units may be measured although the method is most efficient for piece length for which only the count of intersected pieces is required. In line intercept sampling the lengths of interception are measured; the method is suited for estimating the proportion of an area occupied by various cover types. Pielou describes line transect sampling as a means by which the size or density of an animal or bird population is estimated from the sightings made during a linear transversing of the area. Pielou states that "Line transect sampling should not be confused with line intersect sampling or line intercept sampling. These are three entirely different sampling methods". However, as shown by De Vries (1979b), the line transect estimators of the above type can be derived via line intersect theory. The term line transect sampling might also be used to describe sampling by means of a narrow belt transect.

Strictly speaking, the sampling unit is then a rectangular plot, but one which is, in one direction, very long, and in the other direction sufficiently narrow for the items found therein to be uniquely ordered along its length. It would likely be preferable to reserve the term "belt transect" for this situation. By dividing the length into short segments and interpreting the association between adjacent segments, one may be able to make inferences about population structure (e.g. Pielou, 1962, 1963, 1965; Warren, 1972).

Summation

In brief, line intersect sampling is just one more sampling tool. It is not a panacea. Along with nearest-neighbour distances it, perhaps, lies in the category of "indirect methods" which can be tailor made for specific objectives and/or specific situations. When used appropriately these can be highly advantageous and, in general, LIS has proved its worth for the economic assessment of logging residue and the like. Caution is needed when such methods are extended to other objectives; there will often be more efficient alternative methods.

REFERENCES

Bailey,G.R.,1969: An evaluation of the line-intersect method for assessing logging waste. Canadian Dept. Fisheries and Forestry, Information Report VP-X-23, 41pp.

Bailey,G.R.,1970: A simplified method of sampling logging residue. Forestry Chronicle, vol.46 pp.288-294.

Brown,J.K.,1971: A planar intersect technique for sampling fuel volume and surface area. Forest Science, vol.17 pp.96-102.

Brown,J.K. and Roussopoulos,P.J.,1974: Eliminating biases in the planar intersect method for estimating volumes of small fuels. Forest Science, vol.29 pp.350-356.

Buffon,G.,1777: Essai d'arithmétique morale. Supplément à l'Histoire Naturelle, vol.4.

Canfield,R.H.,1941: Application of the line interception method in sampling range vegetation. Journal of Forestry, vol.39 pp.388-394.

De Vries,P.G.,1973: A general theory on line intersect sampling with appliction to logging residue inventory. Mededelingen Landbouwhogeschool Wageningen, Netherlands, No.73-11 23pp.

De Vries,P.G.,1974: Multi-stage line intersect sampling. Forest Science, vol.20 pp.129-133.

De Vries,P.G.,1979a: Line intersect sampling - statistical theory, applications, and suggestions for extended use in ecological inventory. In Statistical Ecology Vol 5: Sampling Biological Populations, R.M.Cormack, G.P.Patil and D.S.Robson (Eds.), International Co-operative Publishing House, Fairland MD, USA, pp.1-709.

De Vries,P.G.,1979b: A generalization of the Hayne-type estimator as an application of line intersect sampling. Biometrics, vol.35 pp.743-748.

De Vries,P.G.,1986: Sampling Theory for Forest Inventory: A Teach-Yourself Course. Springer-Verlag, Berlin. 399pp.

Grosenbaugh,L.R.,1964: Some suggestions for better sample tree measurement. Proceedings of the Society of American Foresters' meeting, 1963 pp. Boston, MA. pp.36-42.

Hájeck,J.,1964: Asymptotic theory of rejective sampling with varying probabilities from a finite population. Annals of Mathematical Statistics, vol. 35 pp.1491-1532.

Hazard,J.W. and Pickford,S.G.,1986: Simulation studies on line intersect sampling of forest residue. Part II. Forest Science, vol.32 pp.447-470.

Howard,J.O. and Ward,F.R.,1972: Measurement of logging residue - alternative applications of the line intersect method. UDSA Forest Service Research Note PNW-163, 8pp.

Martin,A.J. 1976: Suitability of the line intersect method for sampling hardwood logging residues. UDSA Forest Service Research Paper NE-339.

Matérn,B.,1964: A method of estimating the total length of roads by means of a line survey. Studia Forestalia Suecica, vol.18 pp.68-70.

Ménard,M. and Dionne,J.,1972: Essai d'une méthode par transects pour estimer le volume de bois non utilisé lors d'une exploitation forestière. Gouvernement du Québec, Ministère des Terres et Forêts, Service de la Recherche Note no. 2.

Meeuwig,R.O., Miller,E.L. and Budy,J.D.,1978: Estimating pinyon-juniper cordwood with the line-intersect method. USDA Forest Secvice Research Note INT-242.

Pickford,S.G. and Hazard,J.W.,1978: Simulation studies on line intersect sampling of forest residue. Forest Science, vol.24 pp.469-483.

Pielou,E.C.,1962: Runs of one species with respect to another in transects through plant populations. Biometrics, vol.18 pp.579-593.

Pielou,E.C.,1963: Runs of healthy and diseased trees in transects through an infected forest. Biometrics, vol.19 pp.603-614.

Pielou,E.C.,1965: The spread of disease in patchily infected forest stands. Forest Science, vol.11 pp.18-26.

Pielou,E.C., 1985: Line intercept sampling; Line intersect sampling; Line transect sampling. In Encyclopedia of Statistical Sciences, S.Kotz and N.L.Johnson (Eds.) Wiley NY, pp.70-74.

Strand, L.,1958: Sampling for volume along a line. Meddelsen Skogforsoksvesen, Vollebek, Norway, No. 51.

Van Wagner,C.E.,1968: The line intersect method in forest fuel sampling. Forest Science, vol.15 pp.20-26.

Van Wagner,C.E. and Wilson,A.L.,1976: Diameter measurement in the line intersect method. Forest Science vol.22 pp.230-232.

Warren,W.G., 1972: An extension of the belt transect method for assessment of forest disease. Proceedings of the 3rd Conference of the Advisory Group of Forest Statisticians, IUFRO.INRA Publ. 72-3, Paris, France, pp.261-275.

Warren.W.G. and Olsen,P.F.,1964: A line intersect technique for assessing logging waste. Forest Science, vol.10 pp.267-276.

Inventories of U.S. Wetlands

W.E. Frayer
Dean
School of Forestry and Wood Products
Michigan Technological University
Houghton, MI 49931

ABSTRACT

The U.S. Fish and Wildlife Service has a mandate to conduct periodic assessments of wetlands in the United States. The first national report was produced in 1983. The next report to the U.S. Congress is due in 1990. The procedures and results of the inventories performed to prepare the assessments are given in this paper.

INTRODUCTION

Given the best of all possible circumstances, a wetlands inventory would produce complete enumeration of the history, current condition and management alternatives for each portion of all wetlands - as well as for other areas that could be converted into wetland status or that interact with adjacent wetlands. Given funding limitations, an inventory of this type is only possible for small areas - perhaps individual wetlands. Such an inventory is often called a "management inventory" because it provides information needed to manage a specific wetland.

If a wetlands inventory is to cover a large area - such as the United States - other tactics must be used. A well-designed, periodically updated inventory would not produce the information essential for management; but, it can produce what is needed to determine if management is in fact necessary. This kind of inventory has been given various descriptive titles; it is sometimes called a "policy inventory".

What kind of inventory is needed? Both are usually necessary. Unless a broad inventory is conducted and the resulting information points out specific needs, there is little chance of getting adequate support for a comprehensive series of management inventories. What can and has been done in the U.S. is to conduct a broad inventory of the nation's wetlands. This was conducted a few years ago by the U.S. Fish and Wildlife Service with the assistance of other agencies and contractors (Frayer, et al., 1983; Tiner, 1984). It consisted of studying approximately 3,500 sample units to determine changes between the mid-1950s and 1970s. The area sampled was less than 0.5 percent of the nation's area - something quite different from a 100 percent sample common for a management inventory. The results are fairly reliable and very useful.

CLASSIFICATION

In general terms, wetlands are lands where saturation with water is the dominant factor determining the nature of soil development and the types of plants and animal communities living in the soil and on its surface. Technically, wetlands are lands transitional between terrestrial and aquatic systems where the water table is usually at or near the surface or the land is covered by shallow water. Wetlands must have one or more of the following three attributes: 1) at least periodically, the land supports predominantly hydrophytes; 2) the substrate is predominantly undrained hydric soil, and 3) the substrate is nonsoil and is saturated with water or covered by shallow water at sometime during the growing season of each year. Common terms used to describe various wetlands include marshes, swamps, bogs, small ponds, sloughs, potholes, river overflows, mud flats, and wet meadows.

The classification and categories used are described by Cowardin, et al. (1979). Groupings of categories were made to accommodate 1) the special interests of the study and 2) the detail to which available aerial photography could be interpreted. The kinds of wetlands for which estimates were produced include:

Marine intertidal
Estuarine intertidal nonvegetated
Estuarine intertidal emergent
Estuarine intertidal forested and scrub/shrub
Palustrine unconsolidated shore
Palustrine open water
Other palustrine nonvegetated
Palustrine emergent
Palustrine forested
Palustrine scrub/shrub

Within the classification structure that follows, wetlands and deepwater habitats are grouped according to systems. A system consists of environments of similar hydrological, geomorphological, chemical and biological factors. Each system is further divided following traditional ecological concepts. Although deepwater habitats were included in the study, they are not considered in this paper.

The *Marine System* extends from the outer edge of the continental shelf to the high water of spring tides or to the boundary of other systems as defined later. *Marine Intertidal* includes areas in which the substrate is exposed and flooded by tides, including the associated splash zone.

The *Estuarine System* consists of deepwater tidal habitats and adjacent tidal wetlands which are usually semi-enclosed by land, but have open, partially obstructed, or sporadic access to the open ocean and in which ocean water is at least occasionally diluted by fresh water runoff from the land. Offshore areas with typically estuarine plants and animals, such as mangroves and oysters, are also included. *Estuarine Intertidal* is the portion exposed and flooded by tides, including the associated splash zone. For the purposes of this study, *Estuarine Intertidal* wetlands are shown by the following groups: *Nonvegetated*, *Emergent*, and *Forested* and *Scrub/Shrub*. *Nonvegetated* contains no emergent vegetation but does include vegetation in the form of aquatic beds, while *Emergent* contains primarily those erect, rooted herbaceous plants typically found in wet environments. *Forested* is characterized by the presence of trees, and *Scrub/Shrub* includes areas dominated by shrubs and small or stunted trees.

The *Lacustrine System* includes wetlands and deepwater habitats situated in topographic depressions or dammed river channels. Each area must exceed 20 acres or have depths in excess of two meters or have an active waveformed or bedrock shoreline feature. The *Riverine System* includes wetlands and deepwater habitats contained within a channel. Lacustrine and Riverine wetlands were grouped with and included in Palustrine wetlands in this study.

The *Palustrine System* includes all nontidal wetlands not included within any of the other four systems. There are no deepwater habitats included. For this study, *Palustrine* wetlands are shown by the following groups: *Unconsolidated Shore*, *Open Water*, *Other Palustrine Nonvegetated*, *Emergent*, *Forested* and *Scrub/Shrub*. *Unconsolidated Shore* includes wetlands generally having unstable substrates with less than 75 percent cover of stones, boulders, or bedrock and little or no vegetation. *Open Water* includes small inland open water bodies that are too small to be part of the *Lacustrine System*. *Other Palustrine Nonvegetated* includes other inland wetlands with little or no vegetation other than aquatic beds, and the remaining terms are defined as they were under the *Estuarine System*.

In addition to these, three more categories were used in the study. Two are *Urban* and *Agriculture*; and, together with *Other Surface Area* (forests, rangeland, etc., not qualifying as wetland), they account for all other areas.

This is only a brief discussion of the classification used in the study. It is difficult to differentiate the categories further without introducing highly technical terms. For those interested in detailed, exact definitions, the descriptions presented by Cowardin, et al. (1979) are available.

SURVEY PROCEDURE

The objective of the study was to develop statistical estimates of acreage for categories of wetlands for the lower 48 states during the 1950's, the 1970's and the change for the period. A survey was designed to develop national statistics for the 1970's that will, on the average, have a probability of 90 percent that estimated totals are within 10 percent of the true totals, by category.

The sampling design and data compilation procedures were developed to generate flyway and state estimates also. Although these estimates are less reliable, they provide a basis for designing and intensifying flyway or state studies to obtain precise estimates.

A stratified random sample was used with the basic strata being formed by state boundaries and the 35 physical subdivisions described by E.H. Hammond (1970). Additional strata specific to the study are special coastal strata encompassing the Marine Intertidal category and Estuarine System and other strata encompassing the Great Lakes. This resulted in over 200 strata for the study.

Sample units were allocated to strata in proportion to the expected amount of wetland and deepwater habitat acreage as estimated by earlier work (including Shaw and Fredine, 1956). A total of 3635 sample units were used in the study.

Each sample unit is a four-square mile area, two miles on each side. After the units were selected at random within strata and plotted on U.S. Geological Survey topographic maps, aerial photography was obtained for the 1950's and 1970's. The majority of the 1950's photography was 1:20,000 scale black and white and for the 1970's it was 1:40,000 black and white. Scales were adjusted using stereo zoom transfer scopes. The units were photointerpreted in entirety for the 1950's and the changes were photointerpreted on the 1970's photos. All wetland and deepwater habitat changes were marked as to cause, either natural or human induced. The photointerpreted data from each unit were then prepared for computer analysis. Several quality control checks were routinely made to eliminate errors. Photointerpretation and data compilation were completed in July, 1982.

RESULTS

Estimates for the 1950's, 1970's and change during the period were produced for the kinds of wetlands described and are presented in Table 1. Some of the significant results are described below by category.

Changes in Estuarine Nonvegetated Wetlands

The net change in Estuarine Nonvegetated wetlands was small, with an overall net gain. The only change of significance was a loss of 21 thousand acres to urban development. Texas, in the Central Flyway, accounted for almost 10 thousand acres of loss and Florida, in the Atlantic Flyway, over nine thousand acres.

Table 1. Area, in thousands of acres, by kind of wetland for the conterminous United States [1].

	1950's	1970's	Change
Marine Intertidal	82.4 (13.3)	78.4 (14.0)	-4.0 (57.5)
Estuarine Intertidal Nonvegetated	741.1 (9.9)	746.5 (9.8)	+5.4 (>100)
Estuarineal Intertidal Emergent	4,276.0 (4.2)	3,922.8 (4.3)	-353.2 (8.3)
Estuarine Intertidal Forested & Scrub/Shrub	592.1 (13.6)	573.0 (14.4)	-19.1 (93.2)
Palustrine Unconsolidated Shore	293.6 (23.0)	445.6 (33.2)	+152.0 (5.5)
Palustrine Open Water	2,319.9 (6.5)	4,393.1 (7.7)	+2,073.2 (2.5)
Other Palustrine Nonvegetated	90.9 (25.2)	131.8 (23.4)	+40.9 (39.9)
Palustrine Emergent	33,112.6 (15.7)	28,441.4 (17.5)	-4,671.2 (5.2)
Palustrine Forested	55,707.4 (3.7)	49,713.4 (3.6)	-5,994.0 (3.7)
Palustrine Scrub/Shrub	10,998.2 (11.0)	10,611.1 (12.5)	-387.1 (56.7)
All Wetlands	108,214.2 (6.0)	99,057.1 (6.4)	-9,157.1 (3.1)

---

[1] Entries in parentheses are sampling errors in percent (standard error as a percentage of the estimated total).

Changes in Estuarine Vegetated Wetlands

The change in Estuarine Vegetated wetlands resulted in a net loss of 372 thousand acres. Most of this loss (353 thousand acres) was in Estuarine Intertidal Emergent wetlands (coastal salt marshes). Most of the loss was to Estuarine Subtidal deepwater habitats (bay bottoms).

Most of the remaining loss was to urban development which accounted for over 106 thousand acres (the size of a square area almost 13 miles on each side). Two-thirds of the loss was from Estuarine Intertidal Emergent wetlands with the remainder from Estuarine Intertidal Forested and Scrub/Shrub wetlands. The majority of this change occurred in the Atlantic Flyway with Florida accounting for 43 thousand acres. Louisiana, in the Mississippi Flyway, lost 34 thousand acres.

Some Estuarine Intertidal Emergent wetlands changed to Estuarine Intertidal Forested and Scrub/Shrub wetlands. The net change was 21 thousand acres of which 18 thousand acres are in Florida. Some additional Estuarine Intertidal Emergent wetland changes were due to shifts to Estuarine Nonvegetated wetlands. The net change was 19 thousand acres.

Changes in Palustrine Nonvegetated Wetlands

The overall net change in Palustrine Nonvegetated wetlands was a gain of 2.3 million acres.

Changes occurred between Palustrine Nonvegetated wetlands and Palustrine Vegetated wetlands. These changes balanced out for the most part.

A significant net gain came from agricultural land. Over 200 thousand acres were gained, mainly in Palustrine Open Water wetlands, due to construction of farm ponds. The vast majority of gains were in the Central and Mississippi flyways.

A large net gain, mainly in Palustrine Open water wetlands, came from lands not originally classified as agriculture or urban. Over 1.7 million acres were gained, mostly due to construction of ponds. Half of these areas were in the Central Flyway.

Another net gain in Palustrine Open Water wetlands came from Palustrine Forested Wetlands--108 thousand acres.

Changes in Palustrine Vegetated Wetlands

The net change in Palustrine Vegetated wetlands was a loss of 11 million acres. Nearly all the loss was due to agriculture. The overall net loss consists of six million acres of Palustrine Forested wetlands (swamps), 4.7 million acres of Palustrine Emergent wetlands (inland marshes and wet meadows), and the remainder from Palustrine Scrub/Shrub wetlands.

Mississippi Flyway losses were dominate in terms of size; in that Flyway, a net loss of 4.5 million acres occurred in Palustrine Forested wetlands. The vast majority is along the lower Mississippi River in Louisiana, Mississippi and Arkansas.

The next largest loss in the Mississippi Flyway was in Minnesota. Dominant losers in the Atlantic Flyway are Florida and North Carolina. Large losses in the Central Flyway occurred in South Dakota, North Dakota, Nebraska, and Texas. The largest loss in the Pacific Flyway was in California. In general, the states along the lower Mississippi River lost acreage from Palustrine Forested wetlands, while losses in other states were predominantly from Palustrine Emergent wetlands.

Net losses to urban development consisted of 367 thousand acres from Palustrine Forested wetlands, 396 thousand acres from Palustrine Emergent wetlands, and 124 thousand acres from Palustrine Scrub/Shrub wetlands. This total, larger than the size of Rhode Island, is concentrated in the Atlantic and Mississippi flyways. The largest loss in the Atlantic Flyway occurred in Florida. Large losses in the Mississippi Flyway took place in Louisiana, Michigan and Minnesota.

There was a net change of 927 thousand acres of Palustrine Emergent wetlands to Palustrine Scrub/Shrub wetlands and a net change of 693 thousand acres of Palustrine Scrub/Shrub wetlands to Palustrine Forested wetlands.

Net change of Palustrine Emergent wetlands to Palustrine Unconsolidated Shore wetlands occurred on 131 thousand acres, of which 124 thousand acres are in the Central Flyway.

CONCLUSION

The results reported are based on a designed study of the wetlands of the lower 48 states. Because the design involved careful measurement of a sample of the nation's surface area, most of the results are meaningful only at the national level or for broad areas. Some of the results, however, are useful at flyway and state levels. Intensification of the samples for selected areas in future studies can provide useful results for those areas.

Some results are very clear--such as huge decreases in wetlands, especially in the lower Mississippi River states of Louisiana, Mississippi and Arkansas. The next largest loss in the Mississippi Flyway was in Minnesota, with losses also occurring in Michigan, Wisconsin, Illinois, and Alabama. Dominant losers in the Atlantic Flyway are Florida and North Carolina with losses also in Georgia, South Carolina, Maryland, New Jersey, and Delaware. Large losses in the Central Flyway occurred in South Dakota, North Dakota, Nebraska, and Texas. The largest loss in the Pacific Flyway was in California.

Other changes are also clear, but involve lesser acreages. Importance of change, however, is not necessarily reflected by area alone. Some of the smaller wetlands--particularly along the coastline of the United States--are extremely important habitats for plant and animal life.

Significant increases occurred in open water areas. The importance of these new habitats to fish and wildlife populations is yet to be determined.

The vast majority of the loss of Estuarine Intertidal Emergent wetlands (coastal salt marshes) occurred in Louisiana. This resulted from a shift to Estuarine Subtidal deepwater habitats (bay bottoms). The remaining loss of Estuarine Intertidal Emergent wetlands was to urban development, mostly in Florida and Louisiana.

Nearly all the loss of Estuarine Intertidal Forested and Scrub/Shrub wetlands occurred in Florida to urban development.

The net loss in Palustrine Vegetated wetlands was 11 million acres; 15 times the size of Rhode Island; twice the size of New Jersey; as large as the combined states of Massachusetts, Connecticut and Rhode Island. Nearly all the net loss was due to agriculture.

The results of this report document major net losses of wetlands and provide insight to where these net losses are taking place. This report does not document the significant reduction of quality of many wetlands. Some of the factors that cause this reduction in quality are: canals and inlets that cause changes in water chemistry due to saltwater intrusion, mosquito ditching along the Atlantic coast, polluted runoff from adjacent uplands or polluted inflow from rivers and streams, urban encroachment, and dissection by transportation corridors.

As the need for agricultural crops continues to grow and urban areas continue to expand, the total acreage of wetlands will continue to decline. Continual monitoring of wetlands and changes in wetlands is needed to provide the information needed for making wise decisions. The results included in this report provide wetland information similar to the forest and range information required by the Forest and Rangeland Renewable Resources Planning Act and to soils information required by the Soil and Water Resource Conservation Act. The results can be updated in the future on the schedule required under those Acts.

REFERENCES

Cowardin, L.M., V. Carter, F.C. Golet, and E.T. LaRoe. 1979. Classification of wetlands and deepwater habitats of the United States. U.S. Fish Wildl. Serv. 103 pp.

Frayer, W.E., T.J. Monahan, D.C. Bowden, and F.A. Graybill. 1983. Status and trends of wetlands and deepwater habitats in the conterminous United States. Dept. Forest and Wood Sci., Colo. State Univ., Fort Collins. 32 pp.

Hammond, E.H. 1970. Physical subdivisions of the United States. In National atlas of the United States. U.S. Geological Survey, Washington, D.C. 417 pp.

Shaw, S.P. and C.G. Fredine. 1956. Wetlands of the United States. U.S. Fish Wildl. Serv., Circ. 39. 67 pp.

Tiner, R.W., Jr. 1984. Wetlands of the United States: current status and recent trends. U.S. Fish Wildl. Serv. 59 pp.

Precise Tree Frequency Estimates in Horizontal and Vertical Point Sampling

Thomas W. Beers
Professor of Forestry
Department of Forestry & Natural Resources, Purdue University
West Lafayette, Indiana 47907

ABSTRACT

Tree frequency estimates are generally made using class mid-point diameters or heights. This article reiterates the caution that the quality of the estimate depends on the class width being used, and cites a procedure and derived formula to calculate refined (precise) estimates. A formula is given also to calculate the error involved and it is recommended that the precise approach be considered for diameter classes two inches or wider (horizontal point sampling) or for height classes four feet or wider (vertical point sampling).

For the estimation of tree frequencies (number of trees per acre or hectare) using point and/or line sampling, the accepted approach is to employ the fundamental formula for tree factor (Husch et al., 1982). For example, using horizontal point sampling, the number of trees per acre represented by each sample tree (at one point) is symbolized by F_t and can be found by using the formula:

$$F_{t_i} = \frac{F}{BA_i} = \frac{F}{\pi * D_i^2/576} = \frac{576*F}{\pi * D_i^2} \qquad [1]$$

where

F = basal area factor, the constant contribution per acre of each tree to the basal area estimate (per acre, of course),

BA_i = the basal area in sq. ft. for the i^{th} tree, for which F_{t_i} is being calculated,

and D_i = tree diameter in inches.

(With appropriate unit changes, the metric analogy to this formula is obtained by replacing the constant 576 by 40,000).

In the application of this formula, it has been common practice to substitute tree diameter midpoints for the class width appropriate for the level of precision used in the inventory. That is, if precise field measurements are available, say to the nearest tenth-inch, one could use class midpoints like 4.1, 4.2, 4.3, etc., to substitute into formula [1]. On the other hand, data grouped into two-inch classes would demand midpoint diameters like 2, 4, 6, etc. It is rarely recognized that, as Bitterlich (1984, p. 26) stated "The accuracy of the estimate [of tree frequencies per unit area] depends on the narrowness of the diameter class." The implication of this innocent sounding caution is that if classes wider than two inches are used, the estimation error can be considerable, especially for the smaller sizes. The error is caused by the fact that the relation of tree factor to tree diameter is curvilinear, and to use a class midpoint (usually the arithmetic mean of the class boundaries) to represent the class inherently involves a "grouping error."

How should one proceed to avoid such grouping errors? The procedure for developing an appropriate formula has been shown by Bitterlich (1984, Appendix 2) and modified by Beers (1989). Using the referenced procedure, the simplest formula for estimating what can be called a precise tree factor, F_{tp}, for a specified class, is:

$$F_{tp} = \frac{576*F}{\pi} * \frac{1}{D_l*D_u} \qquad [2]$$

where

F = basal area factor,

D_l = lower boundary of the diameter class,

and D_u = upper boundary of the diameter class.

To demonstrate the effect of this refined approach, if two-inch classes are used, the number of trees per acre represented by a 4-inch tree (assuming basal area factor = F = 10) would be either

(1) using the midpoint diameter, 4, and formula [1]:

$$F_t = \frac{576*10}{\pi*4^2} = 114.5916$$

or (2) using the class boundaries, $D_l = 3$ and $D_u = 5$, and formula [2]:

$$F_{tp} = \frac{576*10}{\pi*3*5} = 122.2310$$

The error involved, then, by using the midpoint approximation of 114.59 rather than the precise value of 122.23 is 7.64 trees per acre (at one location, of course), or 6.25 percent of the precise estimate.

In a separate article (Beers, 1989) shows that the error involved by using formula [1] rather than [2] is independent of basal area factor and is a function of class width, w, and midpoint diameter, D_m:

$$E\% = 25*(w/D_m)^2 \qquad [3]$$

where

$E\%$ = the difference in estimates of tree frequency using formula [1] and formula [2], expressed as a percent of the formula [2] estimate.

Applying formula [3] to the case where two-inch classes are used, errors for the 2- and 4-inch classes would be $25*(2/2)^2 = 25$ percent and $25*(2/4)^2 = 6.25$ percent respectively. Similarly, if four-inch classes are used the errors for the 4- and 8-inch classes would be 25 and 6.25 percent.

An illuminating refinement of the precise approach was suggested by Iles[1]. He suggests that formula [2] be re-written as

$$F_{tp} = \frac{576*F}{\pi*(GM)^2} \qquad [4]$$

where

GM = geometric mean of the class boundaries

$= (D_l*D_u)^{1/2}$

This observation points up the difference between the normal and precise approaches, since the normal approach makes use of the squared midpoint diameter, i.e., the squared arithmetic mean of the class boundaries. The geometric mean of the class boundaries is required because the relation between F_t and diameter is not linear; use of the arithmetic mean of the class boundaries in this situation leads to a "two large" diameter and therefore an underestimate of the number of trees per acre.

A legitimate question the practitioner may ask is--how will this estimation error affect my inventory processing? In light of the small errors encountered for narrow diameter classes[2], the usual procedure for calculating the tree factor, formula [1], can be used confidently if tree diameter are recorded to the nearest inch or less. However, for two-inch classes and wider, trees per acre should be calculated using the precise approach, formula [2].

For the estimation of stand characteristics other than tree frequency, the recommended procedure has yet to be developed. The best procedure would be simply to avoid the grouping error by recording field data using precise tree diameters. If wide diameter classes must be used, several approaches are possible, including developing "precise" formulas (analogous to formula [2]) for each basic "factor" such as volume factor, the per acre contribution of each sample tree. Realistically, however, since the usual estimate of the tree characteristic (tree volume, in this case) is relatively imprecise, developing such refined formulas is probably not warranted. A less basic but simpler approach may be to use the precise tree factor formula [2] and obtain the "other" factors as outlined by Husch et al. (1982). That is, for the case of estimating volume per acre, the volume factor (each tree's per acre contribution at one location) would be

$$F_{vp} = V*F_{tp}$$

where F_{vp} = precise volume factor,

V = tree volume,

and F_{tp} = precise tree factor from formula [2].

The inventory planner may find Table 1 useful to obtain the precise tree factors for whatever class width is being used. The table represents a solution of formula [2], but note that if $D_u = D_l$, formula [2] reduces to formula [1]. Therefore the "normal" tree factors are found on the main diagonal. Examples employing the upper and lower halves of Table 1, are shown below.

---

[1] Personal communication, 1988.

[2] The error for the 5-inch one-inch class is 1 percent, for the 5.0 inch tenth-inch class is 0.01 percent, and for the 5.00 inch hundredth-inch class is .0001 percent.

Table 1. Precise Tree Factors for Lower and Upper Diameter Class Boundaries in Inches. Horizontal Point Sampling with F = 10.00. (See footnote for calculation formulas).

Lower	Upper Diameter												
Diameter	2	3	4	5	6	7	8	9	10	11	12	13	14
1	916.73	611.15	458.37	366.69	305.58	261.92	229.18	203.72	183.35	166.68	152.79	141.04	130.96
2	458.37	305.58	229.18	183.35	152.79	130.96	114.59	101.86	91.67	83.34	76.39	70.52	65.48
3		203.72	152.79	122.23	101.86	87.31	76.39	67.91	61.12	55.56	50.93	47.01	43.65
4			114.59	91.67	76.39	65.48	57.30	50.93	45.84	41.67	38.20	35.26	32.74
5				73.34	61.12	52.38	45.84	40.74	36.67	33.34	30.56	28.21	26.19
6					50.93	43.65	38.20	33.95	30.56	27.78	25.46	23.51	21.83
7						37.42	32.74	29.10	26.19	23.81	21.83	20.15	18.71
8							28.65	25.46	22.92	20.83	19.10	17.63	16.37
9								22.64	20.37	18.52	16.98	15.67	14.55
10									18.33	16.67	15.28	14.10	13.10
11										15.15	13.89	12.82	11.91
12	Note: when D(l)=D(u)=1										12.73	11.75	10.91
13		F(t) = 1833.5										10.85	10.07
14													9.35

Lower	Upper Diameter												
Diameter	1.5	2.5	3.5	4.5	5.5	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5
0.5	2444.6	1466.8	1047.7	814.87	666.71	564.14	488.92	431.40	385.99	349.23	318.86	293.35	271.62
1.5	814.87	488.92	349.23	271.62	222.24	188.05	162.97	143.80	128.66	116.41	106.29	97.78	90.54
2.5		293.35	209.54	162.97	133.34	112.83	97.78	86.28	77.20	69.85	63.77	58.67	54.32
3.5			149.67	116.41	95.24	80.59	69.85	61.63	55.14	49.89	45.55	41.91	38.80
4.5				90.54	74.08	62.68	54.32	47.93	42.89	38.80	35.43	32.59	30.18
5.5					60.61	51.29	44.45	39.22	35.09	31.75	28.99	26.67	24.69
6.5						43.40	37.61	33.18	29.69	26.86	24.53	22.57	20.89
7.5							32.59	28.76	25.73	23.28	21.26	19.56	18.11
8.5								25.38	22.71	20.54	18.76	17.26	15.98
9.5									20.32	18.38	16.78	15.44	14.30
10.5										16.63	15.18	13.97	12.93
11.5	Note: when D(l)=D(u)=.5										13.86	12.75	11.81
12.5		F(t) = 7333.9										11.73	10.86
13.5													10.06

Precise factors (F_{tp}) = $(576*F)/(\pi*D_l*D_u)$

Normal factors (F_t), on the diagonal, use the same formula, with $D_l = D_u$.

Example 1. Four-inch classes with midpoints of 4, 8, 12, etc., using the upper half of Table 1.

Class midpoint	Class boundaries D_l to D_u	Precise tree factor F_{tp}
*	0 to 2	*
4	2 to 6	152.79
8	6 to 10	30.56
12	10 to 14	13.10

* For the lower, usually truncated, class the use of the normal tree factor corresponding to the upper class boundary is recommended. Here, from Table 1, the F_t corresponding to $D_u = 2$ is 458.37.

Example 2. Three-inch classes with midpoints of 3, 6, 9, etc., using lower half of Table 1.

Class midpoint	Class boundaries D_l to D_u	Precise tree factor F_{tp}
*	0 to 1.5	*
3	1.5 to 4.5	271.62
6	4.5 to 7.5	54.32
9	7.5 to 10.5	23.28
12	10.5 to 13.5	12.93

* As in the Example 1 footnote, the factor for this truncated class is that for $D_u = 1.5$, i.e., 814.87.

Special note should be made of the fact that Table 1 can be adapted for basal area factors other than 10.0, since the underlying formula is linear with regard to F. Thus, if F = 37.5, precise tree factors would be 37.5/10 = 3.75 times those in Table 1.

As a final caution for users of vertical point sampling, which shows promise for regeneration inventories, the impact of the grouping error is likely to be more pronounced than for horizontal point sampling discussed in the present article. While the formula for obtaining the precise tree factor, Z_{tp} is analogous to formula [2],

$$Z_{tp} = \frac{Z}{H_l * H_u} \qquad [5]$$

where

Z = quadratic height factor,

H_l = lower boundary of the height class,

and H_u = upper boundary of the height class,

the width of classes commonly used may be larger and/or start lower than for diameter classes. For example, if 5-foot classes are used, the underestimate of tree frequencies (using formula [3]) for the 5-foot and 10-foot classes would be 25 percent and 6.25 percent. And, if 2-foot classes were used the error for the 2-foot and 4-foot classes would still be 25 and 6.25 percent, respectively.

Although it is beyond the intended scope of this article, for completeness it is perhaps helpful to close with a comment regarding the tree frequency estimation error in either fixed-size plot sampling or in line sampling.

In fixed-size plot sampling, there is no error caused by the grouping of tree, because the "tree factor" is always a constant: 1/(plot area in acres). In both horizontal and vertical line sampling, however, the problem caused by grouping is similar to that discussed herein for point sampling, but is not nearly as severe. The reader is referred to Beers (1989) for a more complete discussion.

REFERENCES

Bitterlich, W., 1984: The relascope idea. Commonwealth Agricultural Bureaux, England. 242 p.

Husch, B., Miller, C. I. and Beers, T. W., 1982: Forest mensuration, 3rd. ed. Wiley & Sons, New York. 402 p.

Beers, T. W., 1989: Resolution of Certain Grouping Errors in Point and Line Sampling. Purdue Agr. Exper. Sta., Dept. paper, unnumbered, approx. 20 p.

A Comprehensive Concept for Future Forest Inventories

Prof. Dr. Walter Bitterlich
A-5020 Salzburg
Rennbahnstrasse 4 a
Austria

ABSTRACT

The future questions will not be "Permanent plots versus point sampling", or "Why and how critical height sampling?", or "What advantages will Zöhrer's suggestions about sampling with programmed probabilities bring, and how to perform them?", or even the point sampling methods by Hirata and Minowa using vertical angles. All these items are not real "problems" for measurement techniques in the forests, but can be answered by computers alone if we use a comprehensive concept for point sampling, shortly outlined in this article. The basic feature of it is to develop a special science and technology for getting perfectly adapted taper functions for every sample tree in question. More than ever before, the Tele-Relaskop and Spiegel-Relaskop (Wide Scale or CP-Scale) with the new micro attachment can be used for the development of taper functions as well as for controlling them occasionally and/or permanently.

INTRODUCTION

1. The still rapidly increasing capability of computer techniques opens unforeseen possibilities for forest inventories. To take advantage of this progress, the future measurements in the forest should be made on permanent sample points marked by strong capped plastic tubes, sunk in the ground, being used to support an aluminum staff carrying the instrument on its tilting head.

2. The exact horizontal distance from the sample point to every possible tree of a whole sweep has to be determined. For future requirements on the sample, the angle between the north direction and the tree axis (turning right) should also be measured.

3. For each potential sample tree a very close adjusted taper function has to be found. This can be achieved by exact measurements using the Tele-Relaskop and/or the Spiegel-Relaskop following the author's proposals given in his book "The Relascope Idea" (p. 112-115) and in his special programs for the HP-97.

4. Under these basic conditions, any measurement taken, on any kind of "point sample" can be handled by a big computer, e.g.: angle count sampling at breast height or at other heights using different basal area factors, measurement data from fixed-radius plots of various size, critical height sampling with different basal area factors, point sampling with programmed probability by Zöhrer, and even the sampling methods by Hirata or Minowa using vertical angles from one point (see all these methods in the author's book "The Relascope Idea").

5. In the present symposium about "STATE -OF-THE-ART METHODOLOGY OF FOREST INVENTORY" held July 30-August 5, 1989, in Syracuse, New York, USA, the following items are being discussed:

- Sampling with partial replacement
1) • Updating methods for forest inventory
2) • Importance sampling
3) • Critical height techniques
4) • Multi-stage and multi-phase sampling designs
5) • Point-model based sampling
6) • Integrating multiple value forest surveys into timber surveys
7) • Error of inventory biomass or volume estimates
8) • Expressing diameter distribution by mathematical functions
9) • Estimation of stand ingrowth

10) • Estimation of tree and stand mortality
11) • Error of volume or biomass regression
• Line intersect sampling
12) • Spatial analysis of inventory data
13) • Objectives of forest inventory
14) • Forest inventory and growth and yield models
15) • Permanent plots versus point sampling
• Remote sensing and its use in forest inventory
• New regression concepts
• Geographic information systems

The author is convinced that, if the suggestions presented here are used properly, the topics marked by the numbers 1) to 15) can be answered or demonstrated mainly by computers using the same measurement data.

THEORY AND APPLICATION

1. The grid for permanent sample points should be planned in space in such a manner that the distance between the points can be diminished step wise to a minimum of, let us say, 200 or 100 m. Although the marking of the points is under ground and therefore invisible, the system of the points should be shown on the forest management maps. Square grids should be preferred, although grids of equilateral triangles could also be used.

2. There is no doubt that the first installation of the permanent sample points is very expensive. There is first the surveying and the fixing of the plastic tubes below the surface. Then, there is the accurate measurement of the horizontal distance to every sample tree (depending on a minimum dbh and on a maximum of the relative distance to dbh). If, in addition, for each of these trees, the angle from north direction to its axis (turning right) is also recorded, we obtain the polar coordinates of each tree, that can be easily converted to rectangular coordinates by computers. Under such basic requirements structural silvicultural problems can also be handled by computers; see e.g. Dr. K. Johann's method to measure thinning needs in stands of conifers (Osterr. Forstzeitung 12/1988). For each sample tree a separate sheet (or a magnetical device for recording) has to be used to record all the measuring data and quality attributes. The sheets have to be filed carefully in the sequence of the tree direction, starting from the north direction (turning right).

The next step is to fit to each single sample tree a suitable taper function. For this, it is not necessary to measure on each single tree one, two or more upper diameters. A better way is to take a subsample of a few trees from each species and social class. For the rest of the sample trees only the dbh and the total height is needed to fit good taper functions.

The method works as follows: For the subsample trees two additional diameters between the dbh and the top of the tree must be measured by a Tele-Relaskop, a Wide-Scale or CP-Spiegel-Relaskop (called "Four point method"). These four points should be connected now by a third degree polynomial equation, so that any interpolation by computers is made possible. As seldom the tree measurements can be taken from the center point marked by the plastic tube in the ground, a separate observation point must be found for many of the trees of the subsample. These secondary points cannot be marked by plastic tubes and, therefore, in all these cases, measurements have to be taken with the help of tripods. Nevertheless, their location should also be marked for repeated control measurements. This can be done easily, if we install the points in the close neighbourhood (e.g. "2 m NE from tree axis 01200") of another sample tree and put plastic or metal stakes in the ground. The height of the instrument should also be recorded, as e.g. "175 cm over the stake head, or 170 cm over ground". All instrumental measurements are taken with some new micro attachments.

The subsample trees are used as patterns for the rest of the sample trees of the same species, similar crown or similar social class. This can be done automatically by computer programs if the height and the size of the two upper diameters of the pattern tree are being reduced (or enlarged) proportional to the change of the total height and of the dbh respectively, and then a new individual taper function is calculated. As an example, in the program for the small computers HP-97 and HP-41 CV, the measuring values of the pattern trees are stored on separate magnetic cards for each tree. For the bigger computers where any amount of storage room is available, many thousands of basic taper functions can be stored and recalled when needed. The same is valid for any other characteristic of the trees, like bent or forked stems, different shapes of crowns, guessed or true branch volume, biomass etc., and last but not least, any kind of classification of the state of health of trees (e.g. the scales by J. Pollanschütz).

Therefore, the goal of the field work is to visit each tree in the sample and to classify it according to a very carefully thought out scheme . For computers, neither now nor in the future it will matter how many data, measured or estimated will be recorded for each tree. The cruiser does not necessarily need to know any details about the later evaluation of the single sample tree, selected after specified requirements, e.g. : whether it will be included in the evaluation or not, and if so, by what weight or by which sampling method. The big advantage of the technology recommended is that numerous methods of evaluation can be used in a joint process; at the same time and using the same set of sampling data. The different methods of evaluation compete with each other and result in a mutual control. They tell us which procedure is the most efficient for a given problem.

3. Precision measurements on stems of good visibility: The one pole metal-shaft, screwed firmly into the plastic tube and equipped with tilting head and micro attachment, guarantees a constant height of the instrument - Tele-Relaskop or Spiegel-Relaskop - for all the necessary measurements. By this procedure, a point in space is fixed, all inclinations to the height of primarily chosen tree cross sections can be located again and the directions to the cross section diameters which are measured remain unchanged. Tree stems favourably situated for the use of the Tele-Relaskop or the Spiegel-Relaskop can in this way deliver better data for taper functions and stem growth than, for example, measurements by calipers from ladders; in this last case the direction of the measurements cannot be strictly adhered to.

4. Periodical remeasurements: the establishment of the observation points and the determination of the horizontal distances (as well as the angles from the north) for each of the earlier recorded and measured trees has been already made before. The field work is therefore limited to the identification of these trees according to their first or latest data sheets. As a rule only the dbh and the total height will be measured anew, whereas the resulting new form of the stem will be calculated by the computer. Other changes on the tree conditions are to be recorded according to a predetermined scheme. Trees lost by harvesting, catastrophic events or natural death must also be recorded.

In a second operation, all new trees which are now in-trees in regard to the predetermined conditions (due to their increased dbh), have to be measured according to the instructions of a first recording and their data sheets are filed between the other data sheets in the correct order of the angles from the north direction.

At last all diameters are remeasured fromt he one pole metal-shaft by the Tele-Relaskop or the Spiegel-Relaskop on those stems of good visibility selected earlier for the observation of their growth.

INSTRUMENTS AND TOOLS

The manufacturer of the Relaskops, "FOB, Postfach 12, A-5023 Salzburg, Austria", is a member of the IUFRO and considers himself, not without some pride, to be a "private research institute". During the last few years, FOB has developed the instruments and tools necessary for the "permanent comprehensive forest inventory" presented in this paper in close co-operation with the author.

1. After many experiments, both components have been worked out in a manner and a dimension which can be recommended for general use. A hole of about 50 cm depth is digged for the plastic tube, which has thick walls and is 38 cm long including the cap. The one pole metal-shaft is tightly screwed into the plastic tube, whereby protecting it by its cone-shaped lid against contamination on the inside and enabling to put it easily vertical into the prepared hole. For the firm anchoring in the ground four stakes (about 30 cm long and 1 cm strong) are driven down through the holes of the flange at the bottom of the plastic tube using a hammer, not too light, and having a handle of about 70 cm. During the filling up of the hole, this hammer is used to consolidate (press in) the soil and to wedge the plastic tube with stones in order to fix it as vertically as possible. This is controlled best by the use of a spirit level

which is at disposal in the support for the horizontal base target (see manual of the Tele-Relaskop, part 2, Fig. 8) The support can be screwed on the upper end of the one pole metal-shaft. Having finished this part of the work, the one pole metal-shaft is now unscrewed and taken out of the firmly fixed plastic tube. Care must be taken so that no soil would fall into the tube. A ring-shaped cup made for this purpose provides a proper protection against it, also on later occasions when digged out again. The plastic tube is closed with an unbreakable cap and soil is put over it.

Some suggestions were already made as to how to procede on stony or rocky ground, when establishing these observation points with the plastic tubes. Generally speaking, one should not hesitate to spend, in some instances, a little more money for our "brother tree" - for example for the transportation of drillers, concrete, etc. sometimes by helicopters.

The one pole metal-shaft is 170 cm long and, for ease of transportation, it may consist of two or more parts which can be screwed together. It is a tube of 30 mm in diameter and its wall made from Duralumin is 5 mm thick. The lower part has the already mentioned cone-shaped lid rigidly connected and in a distance fitting to the length of the plastic tube. This cone is pressed against the upper edge of the tube, if the one pole metal-shaft is screwed in resulting in an extra strong connection. The lower and the upper part of the one pole metal-shaft has a 10 mm wide cross borehole facilitating the screwing by use of a bolt (for example one of the big stakes mentioned above).

2. Tilting head and micro attachment. Experiments have shown that one pole metal-shafts of the described kind, although they may be constructed in a very heavy way, always show small vibrations during work because of the long arm of the lever from the ground to the instrument. It is also known that up to this time there are no tilting heads with micro attachments, simple and suitable for the field work in the forest, available on the market. For these reasons a micro attachment for precision measurements has been constructed. This new device is very stable and can be used for the Tele-Relaskop as well as for the Spiegel-Relaskops. The value of the Spiegel-Relaskop of the types "Metric-CP" and "Wide-scale" was raised remarkably by this micro attachment, especially for single tree measurements.

3. For the optical measurements of distance see the author's book "The Relascope Idea" (1984), as well as the manuals "Spiegel-Relaskop" and "Tele-Relaskop" obtainable from "FOB, Postfach 12, A-5023 Salzburg, Austria".

REFERENCES

Bitterlich, Walter, 1984: The Relascope Idea. Commonwealth Agricultural Bureaux, Farnham Royal, Slough S1 2 3BN, England, pp.112-115.

Johann, Klaus, 1988: Durchforstung und Forsteinrichtung. Österreichische Forstzeitung, 12/1988, A-1014 Vienna, pp.52-56.

Neumann, Markus & Pollanschütz, Josef, 1988: Taxationshilfe für Kronenzustandserhebungen. Österreichische Forstzeitung, 6/1988, A-1014 Vienna, pp. 27 - 37.

EFFICIENCY OF TWO-PHASE AND TWO-STAGE SAMPLING FOR TREE HEIGHTS IN FOREST INVENTORY

H. Gary Murchison
Associate Professor,
School of Forestry, Lakehead University,
Thunder Bay, Ontario. P7B 5E1

Alan R. Ek,
Professor,
Department Forest Resources, University of Minnesota,
St. Paul, Minnesota. 55108

ABSTRACT

Optimal tree height sampling intensities in forest inventories were investigated using Monte Carlo simulation. Two two-phase and three two-stage rules were compared for statistical and cost efficiency. For the simulations, sampling with replacement was assumed. Diameters and species of all trees were observed on fixed radius plots or horizontal point samples at the first-phase or stage. At the second-phase or stage, heights were observed. Honer's (1967) standard volume functions were used to estimate tree volumes from which stand volumes and variances of these estimates were computed according to each sampling rule simulated. A cost model for two-level sampling is presented and was used to estimate costs. Sample intensities and costs of optimal surveys were predicted.

Sample frames constructed from three stem mapped forest populations located north-west of Sault Ste Marie, Ontario formed the data base (Ek 1969). These are an 8.1 ha aggregated conifer stand, a 16.2 ha randomly distributed northern hardwood stand and a 7.3 ha uniformly spaced pine plantation (Payandeh 1974).

For all simulations, first-level sampling intensity was 50 percent of the sample frame. Simulation of second-level sampling ranged from 10 to 40 percent of the trees observed at the first level.

Absolute bias in total volume estimates ranged from 7.15 to 0.01 percent. Optimal height sampling fractions are two heights per stage-one sample unit and all trees on 7 to 30 percent of phase-one units. The plot based rules were most efficient in terms of precision, but the point sample based rules and the two-stage plot based rule using simple random sampling at both stages were most efficient economically. However, efficiency was clearly affected by tree population spatial patterns.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge support from the USDA, North Central Forest Experiment Station, the Department of Forest Resources, University of Minnesota, St. Paul, MN. and Lakehead University, Thunder Bay, Ontario. C.T. Scott of the USDA, Northeastern Forest Experiment Station, Broomall, PA. contributed data for this project as did the Great Lakes Forestry Centre, Sault Ste. Marie, Ontario.

INTRODUCTION

Forest inventory for volume estimation may be considered a form of sampling in a plane where clusters of elements (trees) forming plots are sampled at a primary level and trees within plots may be sampled at a second level. Plot locations may be randomly chosen with equal probability of selection and trees may be chosen randomly with either equal or unequal probability of selection. Usually forest or stand volume is the variable of interest. Estimates of stand volume are based on estimates of individual tree volumes. Individual tree volumes are estimated using volume tables or functions which use species and diameter at breast height (dbh) and possibly other variables such as height for prediction (Husch et al., 1972; Avery, 1967; Chapman and Meyer, 1949; Bruce and Schumacher, 1942). For example Honer's (1967) standard volume tables are species specific and employ dbh and tree height as prediction variables.

How many tree heights need to be sampled for reliable volume estimation and what form should the sampling take? Information on the number of tree height and diameter samples required to construct volume tables is readily available (Graves, 1907; Bruce and Schumacher, 1942; Chapman and Meyer, 1949; Avery, 1967 and 1975). Almost no information is available for tree height sample fractions needed to apply the volume tables effectively. Loetsch et. al. (1973) state that "only rarely more than 30 heights are sampled per stand" and these are "more or less" subjectively selected leading to possible bias. Common practice calls for the sampling of dominant and codominant trees only, leading to subjective and possibly non-reproducible samples (Husch et al. 1972). This can lead to errors of estimation (Loetsch et al. 1973). Chapman and Meyer (1949) offered a systematic method which does not ensure a non-subjective sample of heights. Ek (1973) states the usual practice is to measure diameters on all trees contained on plots and take a subsample of heights. Ek et al. (1984) estimated individual heights using regional tree height functions based on basal area, dbh and site index. This method is expected to reduce the number of tree height samples required and lead to unbiased results (Murchison, 1984). Bonner (1974) suggested that a random selection of a few tree heights for measurement in an area with estimation of the rest would not affect volume estimates greatly.

Why not measure heights for all trees sampled? Sampling intensity depends on three factors; desired precision, variability of the sample units in question and the cost of sampling (Cochran, 1977). An attribute which is costly to measure and contributes no more information than another less costly item will be sampled less. This is the case for tree height compared to dbh. Also coefficients of variation for tree heights rarely exceed 15 percent, leading to low required sampling intensities for this attribute (Loetsch et. al., 1973). This leads to two-level sampling. Height samples need not be limited to only those trees included on the sample units (Cunia, 1965). At present, few guidelines for the selection of tree height sample fractions exist for two-level sampling in forest inventory.

Our objectives are to list sampling rules which employ subsampling for heights, determine appropriate levels of tree height sampling and compare the sampling rules based on statistical and cost performance.

BACKGROUND AND METHODOLOGY

Sampling variance expressions may be either exact or approximate (Cochran, 1977; Sukhatme and Sukhatme, 1970; Hansen et. al., 1953). For sampling rules where the expressions are exact, sampling variance and thus sampling rule performance can be determined given knowledge of certain population parameters. However, when variance expressions are lacking and cost functions are complicated, some type of simulation comparisons are appropriate. In this study, simulations of forest sampling were carried out on small but real forest populations.

The relative precision of one survey design to a second design is the sampling variance of the first divided by the sampling variance of the second. Relative variance (rel-variance) may be used in place of sampling variance for this purpose (Scott, 1979). Rel-variance is the square of the coefficient of variation (Sukhatme and Sukhatme, 1970). When costs are ignored, relative efficiency may be defined by the relative precision (Scott, 1979). When costs are included, the relative precision resulting for identical survey costs may be used to compare designs. Conversely, survey costs incurred to meet a specific maximum sampling variance may be compared.

Sampling Simulations

In this study, sampling of a single forest stand was assumed. Sample frames consisting of fixed radius plots or points were constructed for each stand (see Table 2). Two-level field sampling was simulated. At the first level, all trees were observed for species and dbh. At the second-level, trees were observed for species, dbh and height. Measurement costs of individual trees and travel within and between sample plots were recorded. Heights measured at the second level were combined with dbh measurements using Honer's (1967) standard volume functions to derive individual tree volumes. These tree volumes were then used to estimate stand volumes and the sampling variance of these estimates. We assumed that tree volumes were error free and that all non-sampling errors were negligible.

Each plot and point test population, see Table 2, was repeatedly randomly sampled up to 2000 times by each of the plot and point sample based rules listed below. For these Monte Carlo simulations, first-level sample intensities were set at 50 percent of the sample frame to produce stable stand volume and variance estimates. Second-level sample intensities ranged from 10 to 40 percent of first-level samples or a minimum of two units. A minimum of two subsamples are required for estimation of sampling variance using the rules. This would be normal practice. Monte Carlo estimates of sampling variance and survey costs were used to predict sample sizes and costs of optimal surveys.

Sample Selection Rules

Sample Rule 1: Two-phase Sampling Using Fixed Area Plots. Both Cochran (1977) and Sukhatme and Sukhatme (1970) provide estimators for two-phase sampling with equal probability of selection at both phases. The first phase consists of a simple random selection of fixed area plots. A simple random subsample of plots forms the second-phase sample. A ratio-of-means estimator (Freese, 1962) is then used to combine the first-phase and second-phase samples to derive total volume estimates. Estimators for mean plot volume and sampling variance estimators used here may be found in Cochran (1977) and Frayer et al. (1979).

Sample Rule 2: Two-phase Point Samples. Johnson (1961, 1965) defined estimators for stand volume on a per unit area basis and the variance of this estimate for double sampling using horizontal points. He assumed first-phase sample points are selected randomly or systematically and individual trees at each point are then selected with probability proportional to tree basal area using an angle gauge.

The second-phase is a randomly or systematically selected subset of the first-phase points. Johnson noted that the estimators for this rule are approximations in the case of systematic sampling. Random selection of plots was assumed for this study.

Sample Rule 3: Two-stage Plots With Equal Probability Subsampling of Trees. This is the simplest two-stage rule with a simple random selection of fixed area plots comprising the first stage. At the second-stage, a simple random subsample of the elements (trees) on each plot are selected.

Frayer et al. (1979) provide the estimator (function 5.3.21 on page 40) for mean per sample element. A sample based variance estimator for the variable of interest was also given by Frayer et al. (1979, p. 40, function 5.3.22). This expression is not the sampling variance attributable per sample element, but is an expression of the sampling variance per plot. Murchison (1984) gave a function which estimates the sampling variance per element. To derive the total for the population and its sampling variance the number of sample elements in the population, is used as a multiplier (Frayer et al., 1979, functions (5.3.23) and (5.3.24).

Sample Rule 4: Two-stage Plots With Unequal Probability Subsampling of Trees. At the first-stage, simple random sampling is carried out using fixed area plots. On each plot a cumulative list of tree diameters provides a basis for selection of the second-stage sample trees with probability proportional to basal area.

This method facilitates control over sample sizes at the second-stage more easily than using an angle gauge for selecting trees with probability proportional to basal area. Further it avoids any covariance between sample trees and their location with respect to the plot center.

Murchison (1984) defined mean and sample variance estimators for this rule (p 62, functions [3.39] and [3.43] respectively) assuming sampling with replacement.

Sample Rule 5: Two-stage Point Sampling With Equal Probability Subsampling of Trees. Palley and Horwitz (1961) used point samples as primary units and subsamples of trees from points as secondary units. Schreuder (1966 and 1970) expanded on this procedure giving point and variance estimators for both population parameters and samples. Langley (1975) recommended this sampling scheme as being unbiased and efficient.

Sampling Efficiency and Survey Cost Model

For surveys which employ multiple levels, optimum sampling implies the simultaneous optimization of sample sizes at each level. Sukhatme and Sukhatme (1970, p. 171, function 186) provide an estimate of optimum variance of the variable of interest (Y) based on some fixed cost and estimates of the standard deviation of Y, standard error of predicted Y, and costs of measuring first and second-phase units. These may be obtained from other surveys or a preliminary sample. The optimum second-phase sample size may be taken from Sukhatme and Sukhatme (1970, p. 171, function 185). This expression also gives the optimal size of a first-phase sample. These are functions of the variance of Y, the mean squared error of Y, first and second-level sample unit costs and overhead.

For two-stage sampling, Cochran (1977, p. 281) provides an equality which minimizes variance for a fixed cost or cost for a fixed variance. This leads to an estimate for the minimum direct costs of two-stage sampling for a fixed variance (Cochran 1977, p. 281).

Cochran (1977) approximated the optimal number of second-stage samples for a survey (p. 314, function 11.60) based on knowledge of stage-one and two sample unit costs, stage-one and two variance and the mean number of elements per stage-one unit. This optimum second-stage sample size allows estimation of the direct sampling costs which is used in Cochran's function (1977, p. 280) to estimate the optimal first-stage sample size. These estimates of optimal first and second-level sample sizes may be used in a cost model to estimate survey costs.

For a two-level sample survey cost model, the first and second-level sampling costs should be separate to allow estimation of first and second-level sample sizes. First-level costs are common for both two-phase and two-stage sampling but will vary according to the method of cluster selection; plot sampling or point sampling. Second-level sample units costs are based on plots for two-phase sampling and trees for two-stage sampling. Other costs not directly part of the sampling costs of the two levels include overhead, travel to the forest, presurvey and postsurvey costs. Overhead costs are fixed costs. Any portion of fixed costs attributed to a survey is arbitrary and may be a function of the overall time spent on a survey. Presurvey costs include costs of planning a survey, and acquiring necessary personnel and equipment for the survey. Postsurvey costs include data editing, data processing, reporting and any penalties which may be incurred as a result of the survey.

Four types of travel costs are incurred in the field phases of a forest inventory; (i) travel costs to get to the vicinity of the survey, (ii) travel costs from a vehicle to the field sample units, (iii) travel costs between sample units, and (iv) travel costs within sample units. The first cost includes travel time and vehicle costs for the survey crew to travel from a central field office location to the survey site and return on a daily basis multiplied by the number of days required for field measurements. For remote locations, daily room and board costs for field crews may be incurred and included in this cost item. Since the number of days required is dependent upon sample size, cost (i) is a function of sample size. The last three travel costs are functions of sample unit sizes and sampling intensities at both levels.

A model for the total cost of a two-level survey is:

$$\begin{aligned} TC = {} & Co+2\cdot nd(Cf\cdot df)+2\cdot nc\cdot nd(Cw\cdot dl)+npo\cdot CL1+(npo-nd)\cdot nc\cdot Cw\cdot dp+nso\cdot CP2 \\ & +npo\cdot mo\cdot CS2+npo\cdot mo\cdot Cw\cdot dt+CPAP+Closs \end{aligned} \qquad [1]$$

where, Co = overhead costs assigned to the survey,
Cf = travel cost per unit distance for travel of vehicle and crew from central location to field,
df = distance to field location from central office,
nc = number of people in crew,
nd = number of days required for survey; assumed an integer,
Cw = cost of walking per unit of distance crew,
dl = average distance of first and last sample units from the vehicle,
dp = average distance between primary sample units,
dt = average distance between second-stage sample units,
CL1 = cost of enumerating a first-level sample unit,
CP2 = cost of enumerating a second-phase sample unit,
CS2 = cost of enumerating a second-stage sample unit,
CPAP = cost of planning a survey and processing the results,
Closs = the costs of all penalties incurred as a result of the survey,
npo = the number of phase or stage-one units sampled,
nso = the number of phase-two units sampled,
mo = the number of stage-two units sampled.

For the purposes of this study, Closs, CPAP and Co are assumed constant for all methods and set at zero.

The number of days required for a survey is a function of the time per day available for field work, the time required per unit of sample plus the amount of time spent walking to and from the field and between sample units. The approximate number of days required for a survey may be estimated as:

$$nd = [npo \cdot [TL1+Tw \cdot dp]+nso \cdot TP2+npo \cdot mo \cdot [TS2+Tw \cdot dt]]/[FT-2.0 \cdot Tw \cdot dl] \qquad [2]$$

where, FT = the amount of time per day left for field work after travel to and from the field,
TL1 = time required to measure first-level sample unit,
TP2 = time required to measure second-phase sample unit,
TS2 = time required to measure second-stage sample unit,
Tw = time required to walk per unit distance.

Costs may be expressed in monetary or time units. Since vehicle costs may not be easily expressed in time units, monetary units were used.

The cost of first-level sampling is defined as:

$$CL1 = nc \cdot [Mi\ [Cs+Cw \cdot dt]+Cl] \qquad [3]$$

where, Mi = the number of elements per first-phase or stage sample unit,
Cs = the cost of species identification and dbh measurement per tree,
Cl = the cost to identify the cluster at the selected location.

Two-phase and two-stage sampling are expected to involve different numbers of second-level sample units at each location and require separate cost functions. If we assume that second-phase sampling will be conducted concurrently with the first-phase sample, tree location, identification, dbh measurement and intertree travel costs may be accounted for in the first-phase costs. Second-phase sampling costs per location are given by:

$$CP2 = nc \cdot Mi \cdot Ch \qquad [4]$$

where, Ch = the cost to measure the height of a tree.

Second-stage sample unit costs are defined as:

$$CS2 = nc \cdot [Cs+Ch+Clt+Cw \cdot dt] \qquad [5]$$

where, Clt = the cost to locate a tree for sampling.
Expression [5] assumes that the tally for the second-stage trees is recorded separate from that for the first-stage.

Data Base

The data base for this study consisted of stem mapped data from three forest stands located northwest of Sault Ste. Marie, Ontario and specific costs. Characteristics of these three stands were described by Ek (1969) and Payandeh (1974) and are given in Table 1.

As shown by Pielou's indices in Table 1, the softwood stand shows a strongly aggregated or clumped spatial distribution while the hardwood stand is nearly randomly distributed and the red pine plantation shows a strong uniform spatial distribution.

Table 1; Individual Stand Characteristics

Stand Characteristics	Spruce-fir Stand	Northern Hardwood Stand	Red pine Plantation (unthinned)
Area in hectares	8.09	16.19	7.28
Total number of trees per stand	6811	5503	9066
Pielou's Nonrandomness Index for tree spatial distribution (Payandeh 1974)	1.594	0.975	0.713

Plot and point sampling programs were run in each of the three stands to create test populations of plot and point data for sampling simulation. Characteristics of the resulting six test populations are given in Table 2.

Individual costs assumed for this study were; labor costs of $15.00 per hour per person for a two-person crew. Operating costs for a vehicle were set at $100.00 per day. The average travel distance of field locations from a central office was 75.0 kilometres at an average travel speed of 50.0 kilometres per hour. Walking speed of the crew was set at .2 kilometres per hour for both inter-tree and inter-plot travel. This appears to be low, but allows for significant delays due to frequent stops for taking measurements. Working time per day was assumed to be 480.0 minutes. All costs were held constant for all simulations.

Table 2; Characteristics of Plot and Point Sample Test Populations

Test Population Characteristics	Spruce-fir Stand	Hardwood Stand	Red pine Plantation
(a) Plot Based Populations			
Number of plots in sample frame	242	203	363
Plot size (hectares)	0.025	0.062	0.015
Average number of trees per plot	21.7	20.9	19.3
Stand volume based on all plots (m^3)	1178.88	3171.52	2118.36
Sampling variance of stand volume	2211.24	13080.59	2426.35
(b) Point Based Populations			
Number of points in sample frame	153	138	300
Basal area factor of angle gauge	2.0	3.0	4.0
Average number of trees per point	9.7	7.5	9.2
Stand volume based on all points (m^3)	1131.24	3167.29	2068.62
Sampling variance of stand volume	910.41	10205.94	730.91

For predicting the requirements of optimal surveys, the maximum allowable sampling variance was 25 cubic meters per hectare.

RESULTS

Average Accuracy and Precision of Sample Rules

Average Monte Carlo estimates of stand volumes and variances compared to test population values are given in Table 3. All rules except rule 3, produced differences of less than 0.2 percent of population volumes. Rule 3 produced positive differences of 1.1 to 5.9 percent. We are currently investigating this apparent deviation.

For all rules in all stand conditions, Monte Carlo estimates of variances computed from stand volume estimates varied little between secondary sample fractions. This indicates that the rules performed consistently and supports the averaging of results. Since little difference was found between Monte Carlo variances for different fractions of secondary sampling, the major component of variance of stand volume was due to first-phase or first-stage sampling. This has also been noted by Yandle et. al. (1983) and Scott (1981). As indicated by the relative precision values in column 6 of Table 3, no single rule outperformed the others in all three stand conditions. The point sample rules appear to be the most efficient in a clustered stand, but least precise in random and uniform stands. Under these latter conditions, the two-stage plot based rules are marginally more precise that the two-phase rule using plots. This inconsistency confirms that population characteristics affect sample rule performance.

Efficiency of Sample Selection Rules

If sample intensity is used as a criteria for selecting a sample rule or design, some estimate of efficiency is required. Relative efficiency is defined as the inverse of the ratio, given in percent, of the sample units required by one design divided by the number of sample units required for an alternate design. Relative efficiencies for all 5 rules in the 3 populations are given in Table 4. The relative efficiencies of rules 2, 3, 4, and 5 for both levels of sampling are given in relation to rule 1. Relative efficiencies at the first level are based on the numbers of plots or points required for optimal surveys. At the second level, these are based on the optimal number of trees required. Relative efficiencies based on predicted costs of optimal surveys are also given in Table 4.

Table 3; Accuracy and Precision of Sample Rules in Estimating Total Volumes of Populations.

Population	Sample Rule	Estimated Volume (m^3)	Bias (%) of Volume Estimate	Average Monte Carlo Variance	Relative Precision (%)
Spruce-fir	1	1177	-0.16	4105	100
Stand	2	1132	0.05	2670	142
	3	1249	5.93	4692	98
	4	1179	0.01	3017	136
	5	1132	0.07	1852	205
Northern	1	3172	0.02	11873	100
Hardwood	2	3167	-0.07	21348	55
Stand	3	3206	1.10	7663	158
	4	3171	-0.02	11686	101
	5	3169	0.05	21275	56
Red Pine	1	2119	-0.02	632	100
Plantation	2	2069	0.04	1629	37
	3	2160	1.99	856	78
	4	2119	0.02	628	102
	5	2051	-0.04	1529	40

In terms of the number of primary sample units required, Rule 4 is poorer than the other two-stage rules and about the same as the two-phase rules. Rule 5 appears more efficient than rule 3 for random and uniform conditions, but less efficient in the clustered stand.

The ratings given in column 6 of Table 4 show that Rule 5 is again the most efficient in terms of the number of tree heights to be sampled. Rule 2 is the most efficient for tree height samples in the random and uniform stands but least efficient in the clustered stand. Rule 3 is the most efficient in the clustered stand. Rule 4 appears to be only slightly more efficient than Rule 1, these two rules being much less efficient than the other three rules in terms of the number of tree heights to be sampled.

The effects of combining the optimal sampling intensities at both levels for the various rules are reflected in the relative cost efficiencies in the last column of Table 4. These show the Rule 5 is the most efficient for sampling the uniform and random stands. Rule 2, the other point sample based rule is the second most efficient for the these stands but the least efficient in the clustered stand. In this stand, Rule 5 is the second most efficient. Rule 3 appears to be the most efficient in the clustered stand and of intermediate efficiency in the random and uniform populations. Again, Rules 4 and 1 are the least efficient in overall.

Sampling Intensity

As shown in Table 4, the two-stage rules require only 2 tree heights per stage-one sample unit to satisfy precision requirements. This represents about 10 percent of the trees on stage-one plots and 22 to 29 percent of the trees on stage-one points. For two-phase plots, 7 to 18 percent of the plots must be double sampled. With two-phase points this range shifts to 11 to 28 percent. The higher percentages of height samples required for points reflects the lower number of trees per first-level unit compared to plots.

Table 4; Average requirements of optimal survey as defined by the Monte Carlo Simulations.

Population	Sample Rule	First-Level Samples Required	R.E. (%)	Second-Level Samples Required	R.E. (%)	Level-2 Sample Fraction (%)	Total Survey Cost ($)	R.E. (%)
Spruce	1	299	100	54	100	17.57	32214	100
-fir	2	1008	30	254	48	28.01	66532	48
Stand	3	108	277	2	543	9.52	9985	323
	4	509	59	2	116	9.09	46238	70
	5	295	101	2	197	22.22	184691	74
Northern	1	196	100	13	100	7.05	20060	100
Hardwood	2	102	193	14	259	14.56	56403	56
Stand	3	70	279	2	194	10.00	75082	67
	4	176	111	2	77	9.52	18564	108
	5	65	300	2	209	28.57	43954	56
Red pine	1	93	100	10	100	10.58	8355	100
Plantation	2	58	160	6	345	10.71	3260	256
	3	56	166	2	172	10.53	5012	167
	4	79	118	2	122	10.00	6948	120
	5	37	251	2	261	22.22	2432	344

Note: R.E. is relative efficiency of Rules 2, 3, 4 or 5 versus Rule 1.

DISCUSSION

Using large phase-one or stage-one sample fractions and with replacement sampling for the simulations, repeated sampling of some sample units obviously occurred. This probably reduced estimates of Monte Carlo sampling variance. Thus under operational conditions one could expect all sampling variance estimates to increase, since repeated sampling would be less likely to occur.

How many tree height samples should be recorded during a forest inventory? This study suggests that two tree heights per primary sample unit are required for two-stage sampling and that all trees on 7 to 30 per cent of the primary sample units are required for two-phase sampling. If the above effect of with replacement sampling is correct, these sample intensities may be low. Also under operational conditions, the first-level sampling intensity would be expected to decline. Thus second-level sampling intensities could be expected to increase marginally under operational conditions. However, the primary sample units were small compared to operational plots of 0.04 ha or larger. Increasing the size and therefore the number of elements in plots at the primary level could cause the secondary sample intensity to decline marginally. This area warrants further study.

If the allowable standard error restriction were relaxed, the optimal primary sample levels would be reduced. For two-phase sampling under relaxed limits, optimal phase-two sample sizes could also be expected to decline, but to a lesser extent than for first-phase sample fractions. This is because the major component of sampling variance is between primary sample units rather than within these units. Similarly for two-stage sampling, optimal first-stage sample fractions would be expected to decline if the maximum allowable standard error were increased. However, stage-two sample fractions would not decline as they are already at the minimum of two tree heights per stage-one plot or point.

This study does show that reliable stand estimates may be achieved with small clusters of trees as sample units with subsampling for tree heights. Conventional single level sampling with all tree heights measured also represents costly oversampling. Subsampling, as simulated here, can lead to substantial economies in forest sampling.

LITERATURE CITED

Avery, T.E. 1975. Natural Resources Measurements. 2nd. ed. McGraw-Hill. New York. 318 pp.

----------- 1967. Forest Measurements. McGraw-Hill. New York. 260 pp.

Bonner, G.M. 1974. Estimation versus measurement of tree heights in forest inventories. Forestry Chronicle, Oct. 1974. pp. 200.

Bruce, D.B. and F.X. Schumacher. 1942. Forest Mensuration. McGraw-Hill, New York. 401 pp.

Chapman, H.H. and W.H. Meyer. 1949. Forest Mensuration. McGraw-Hill. New York. 501 pp.

Cochran, W.G. 1977. Sampling techniques. 3rd. edition Wiley Press. New York. 417 pp.

Cunia, T. 1965. Some theory on reliability of volume estimates in a forest inventory sample. For. Sci. 11(1):115-128.

Ek, A.R., E.T. Birdsall and R.J. Spears. 1984. A simple model for estimating total and merchantable tree heights. USDA For. Ser. Res. Note NC-309. 5 pp.

Ek, A.R. 1973. Performance of regression models for tree height estimation with small sample sizes. IUFRO Statistics in forestry research. Proceedings of IUFRO Subject Group S6.02. Vancouver, B.C. pp.67-80.

Ek, A.R. 1969. Stem map data for three forest stands in Northern Ontario. Can. Dept. of Fisheries and Forestry, Can. For. Ser. Info. Rep. O-X-113. Sault Ste. Marie, Ont. 9 pp.

Pelz, D.R. 1976. IFSAS-the Illinois forest sampling simulator. Forestry Research Rep. No. 76-1. Dept. of For. University of Illinois. Urbana. 27 pp.

Schreuder, H.T. 1970. Point sampling theory in the framework of equal-probability cluster sampling. For. Sci. 16(2):240-246.

--------------- 1966. Unequal probability and double sampling in forestry. Doctoral dissertation. Iowa State University, Ames, Iowa. 139 pp.

Scott, C.T. 1981. Simplified estimators for sampling with partial replacement on multiple occasions. Dept. of Forest Resources, Univ. of Minnesota, Staff Series Paper Number 23. St. Paul, Mn. 48 pp.

----------- 1979. Midcycle updating: some practical suggestions. Proceedings of Workshop: Forest Resources Inventories. SAF and IUFRO. Colorado State University, Ft. Collins, Co. pp. 362-370.

Sukhatme, P.V. and B.V. Sukhatme. 1970. Sampling theory and applications. 2nd. ed. Iowa State University Press. Ames, Iowa. 446 pp.

Yandle, D.O., J.R. Myers and H.V. Wiant Jr. 1983. A comparison of some two-stage sampling designs. Renewable Resource Inventories for Monitoring Changes and Needs. Proc. of Int. Conf. Aug. 15-19, 1983. Corvallis, Oregon. pp. 645-6

Frayer, W.E. et al. 1979. Multi-level sampling designs for forest inventories. Dept. of Forest and Wood Sciences, Colorado State University and Rocky Mountain Forest and Range Exp. Stn. USDA For.Ser. CSU Project Report 31-1470-1468. 113 pp.

Freese, F. 1962. Elementary forest sampling. USDA Agr. Handbook No. 232. 91 pp.

Graves, H.S. 1907. Forest Mensuration. Wiley and Sons. New York. 437 pp.

Hansen, M.H., W.N. Hurwitz and W.G. Madow. 1953. Sample survey methods and theory. Vol.1; Methods and applications. Wiley and Sons. New York. 624 pp.

Honer, T.G. 1967. Standard volume tables and merchantable conversion factors for the commercial tree species of central and eastern Canada. Forest Management Research and Services Institute. Info. Rep. FMR-X-5. Can. Dept. of For. and Rural Dev. Ottawa. 153 pp.

Husch, B., C.I. Miller and T.W. Beers. 1972. Forest mensuration. 2nd. ed. Ronald Press. New York. 373 pp.

Johnson, F.A. 1965. Sample sizes for timber cruises. USDA For. Ser. Res. Note PNW-30. 10 pp.

------------- 1961. Standard error of estimated average volume per acre under point sampling when trees are measured for volume on a subsample of all points. USDA For. Ser. Res. Note PNW-201 6 pp.

Langley, P.G. 1975. Multistage variable probability sampling: theory and use in estimating timber resources from space and aircraft photography. Doctoral dissertation. U. of California, Berkely,Ca. 92 pp.

Loetsch, F., F. Zohrer and K.E. Haller. 1973. Forest Inventory. BLV Verlagsgeselleschaft. Munich. Vol.2. 452 pp.

Murchison, H. G. 1984. Efficiency of Multi-phase and Multi-stage sampling for tree heights in forest inventory. Doctoral Dissertation. University of Minnesota, St. Paul. Minnesota. 158 pp.

Palley, M.N. and L.G. Horwitz. 1961. Properties of some random and systematic point sampling estimators. For. Sci. 7(1):52-65.

Payandeh, B. 1974. Spatial pattern of trees in the major forest types of Northern Ontario. Can. Jour. For. Res. 4:8-14.

Payandeh, B. and A.R. Ek. 1971. Observations on spatial distribution and relative precision of systematic sampling. Can. Jour. For. Res. 1:216-222.

Sampling for logging residue and waste in British Columbia

Stephen A. Y. Omule

Ministry of Forests, Research Branch
31 Bastion Square, Victoria BC, Canada V8W 3E7

and

David W. Ormerod

Sylvametrics Consulting
4672 Cordova Bay Road, Victoria BC, Canada V8X 3V7

ABSTRACT

This paper describes a recently implemented sampling design for estimating logging residue and waste in British Columbia, Canada. The estimates are used for billing of stumpage for timber cut from Crown land and for volume-based cut control. *Residue* and *waste* refer to timber that is within the cutting specifications of the harvesting license, but is not removed from the logged area. Residue is composed of pulp or lumber reject grades and does not have to be removed from the logging site. It is part of the license volume. Waste refers to grades of timber that should have been removed from the harvested area. The target population is the volume of residue and waste generated in a calendar year within specified logging areas. Residue and waste are sampled based on stratified single-stage (random) sampling or stratified two-stage sampling. The single-stage design is usually applied to small areas such as single cut-blocks, and the two-stage design is applied to large areas involving several cut-blocks. Stratification criteria include harvesting system and operator, timber type, and cut-block size. In both designs, the sampling units are fixed-area plots systematically located in the logged areas, including landings and other accumulations. Residue and waste on landings may be measured 100% or estimated using 3P sampling, in single-stage sampling. The first-stage sampling units in the two-stage design are cut-blocks selected at random with equal probability. The number of sample plots in the single-stage design or the number of sample cut-blocks in the two-stage design is determined by the desired allowable sampling error and the population coefficients of variation. The number of sample plots for the two-stage design is determined by sampling cost and variability in the two stages. The maximum allowable sampling error varies with the relative amounts of avoidable and unavoidable residue and waste. Residue and waste are classified and measured based on current scaling procedures.

INTRODUCTION

The quantity of residue and waste produced from timber harvesting operations is currently a sensitive issue in the province of British Columbia (B.C.), Canada. Poor market conditions in the recession years resulted in excessive amounts of residue and waste being left on logging sites by operators trying to maintain profitable margins. The public has reacted strongly against continuance of sympathetic policies which allowed such levels of waste to occur.

Credible survey methods had to be developed for classifying and estimating the volume of logging residue and waste generated during timber harvests. The estimates are used for billing of stumpage for timber cut from Crown land and for volume-based cut control. This paper describes a recently implemented survey sampling design used by the Forest Service and major forest companies in B.C. for estimating logging residue and waste.[1]

DEFINITIONS

All tenures for forest harvesting on Crown land in B.C. require a specific cutting authority (usually a cutting permit) unless the timber is sold on the basis of a cruise. The authority allows the tenure holder to remove a specified volume of timber subject to the provisions of the province's Forest Act and the specific harvesting license. Clear-cutting is the common timber harvesting method, and stumpage is normally billed from the scaled volume of harvested wood.

Cutting authority documents for the removal of timber from the Crown lands specify Minimum Utilization Standards (MUS) which determine the merchantable species, log grades, and minimum dimensions for logs that will be applied in measuring the volume of harvested timber.

Residue is composed of log grades that do not have to be utilized, but is part of the cut control volume. Waste refers to wood volume left on the logging area that is within the MUS and is not residue. Residue and waste may be classified either as avoidable or unavoidable. Unavoidable material is the component of residue or waste which cannot be removed with a reasonable and safe effort because of physical impediments. The balance of the residue and waste material is avoidable. Residue and waste surveys will be referred to simply as residue surveys.

SAMPLING OBJECTIVE

The sampling objective of residue surveys is to estimate the volume of residue and waste in a logged area to a given level of precision that is consistently reasonable for stumpage billing and cut control. The statistical precision is expressed as a percentage sampling error at a given probability level and varies with the relative amounts of residue and waste according to the following relationship:

$$E = 25 - [24A_w + 10(A_r + U)]/(S + A_w + A_r + U) \qquad (1)$$

where E = percent sampling error at the 95% probability level
A_w = avoidable waste volume
A_r = avoidable residue volume
U = unavoidable residue and waste volume
and S = scaled volume

Note that the sum of S, A_w, A_r, and U approximates the volume utilization limits of the licence, as defined by its MUS.

---

[1] The paper is based on the sampling design outlined in the manual " Residue and waste measurement procedures for British Columbia", 1989, by B.C. Forest Service, Ministry of Forests, Valuation Branch, 1450 Government Street, Victoria, B.C. Canada. V8W 3E7.

POPULATIONS

The population of interest in residue surveys is the volume of residue and waste available for measurement within a calendar year within a specified cut-block (Cut-block Population) or aggregation of cut-blocks (Aggregated Population). An Aggregated Population is generally much larger than a Cut-block Population.

Residue and waste are distributed across the logged areas in dispersions or accumulations. Accumulations occur at landings and other areas where logs have been yarded or skidded prior to loading. Dispersions occur in the rest of the cut-block. Accumulations and dispersions are sampled as separate sub-populations to avoid the complexity that would arise if they were treated as strata. The sub-populations are generally highly variable, typically with coefficients of variation of 20% to 90%.

For both populations, the residue and waste volumes in the sub-populations are estimated based on stratified random sampling, stratified two-stage random sampling, and 3P sampling. These methods are now briefly described.

CUT-BLOCK POPULATION

After harvesting is completed, areas to be surveyed are delineated and stratified based on timber type, and harvesting system and operator. Dispersed and accumulated sub-populations are distinguished. Dispersed sub-populations are sampled using stratified (single-stage) random sampling. Accumulations are either measured 100% or are sampled based on stratified random sampling and/or 3P sampling.

Stratified Random Sampling

In stratified random sampling, fixed area circular or rectangular plots are systematically established throughout the area on a square grid. Plot size varies with the sub-population: the plot size is 0.04 ha or larger in the dispersions and 0.005 ha or larger in the accumulations. Note that plot size is uniform within a given cut-block. The grid spacing is determined by the sub-population area and number of plots. Rectangular grids may be used in accumulations.

The number of required sample plots, n, for a given sub-population is

$$n = \sum_{j=1}^{l} N_j t^2 C^2 / (\sum_{\mathrm{j}=1}^{\mathrm{l}} \mathrm{N_j E^2 + t^2 C^2}) \qquad (2)$$

where C = sub-population coefficient of variation weighted by strata areas.
E = allowable percent sampling error as defined in equation 1.
N_j= total possible number of plots in stratum j.
t = t-distribution value at the 95% probability level with $n - 1$ degrees of freedom.
and l = total number of strata in the population

The n sample plots are allocated to the l strata in proportion to strata areas. That is

$$n_j = n \frac{H_j}{\sum_{j=1}^{l} H_j} \tag{3}$$

where n_j = number of sample plots in stratum j
and H_j = size of stratum j (ha)

Dimensions of wood pieces within plot boundaries are measured based on log scaling procedures[2] to enable calculation and classification of residue and waste volume. Residue and waste volumes are classified as avoidable (A_r or A_w, respectively) or unavoidable (U). Dimensions and classification of individual pieces that are buried or are inaccessible are visually estimated. In deep accumulations, a visual estimate is made of the volume of the residue and waste that cannot be directly measured. Standing or down trees and snags are measured as if they were logs.

Estimates of W_{ut} (m^3/ha), the mean volume of residue and waste (A_r, A_w, and U) of type u in sub-population t (dispersed or accumulated), and associated confidence limits, are

$$\hat{W}_{ut} = \sum_{j=1}^{l} H_j \bar{w}_j / \sum_{j=1}^{l} H_j \pm t \sqrt{(s^2 / \sum_{j=1}^{l} n_j)(1 - \sum_{j=1}^{l} n_j / \sum_{j=1}^{l} N_j)} \tag{4}$$

where H_j = size of stratum j (ha)
w_{ij} = volume of residue and waste in plot i in stratum j (m^3/ha)
$\bar{w}_j = \sum_{i=1}^{n_j} w_{ij}/n_j$
n_j = number of sample plots in stratum j
s = weighted standard deviation of plot waste and residue $\{w_{ij}\}$ for all strata
$= \sqrt{\sum_{j=1}^{l} H_j^2 s_j^2 / (\sum_{j=1}^{l} H_j)^2}$
s_j = standard deviation of plot waste and residue $\{w_{ij}\}$ in stratum j
$= \sqrt{\sum_{i=1}^{n_j} [w_{ij} - (\sum_{i=1}^{n_j} w_{ij}/n_j)]^2 / (n_j - 1)}$
t = t-distribution value at the 95% probability level
with $n - l$ degrees of freedom.
and all other symbols are as previously defined.

3P Sampling

The 3P sampling method is applied to accumulated residue and waste sub-populations by estimating the residue and waste volumes in all accumulations and measuring all the residue and waste in each of a random sample of accumulations (units). Large accumulations may be sub-divided into smaller units. The application of 3P sampling in residue surveys differs somewhat from the classical definition in that a complete list is made by visually estimating the volumes of all units, so that the largest unit and the total estimate are known. Procedures for measurement and classification of individual pieces of wood are the same as those outlined above.

Estimates of the mean volume of residue and waste of type u in sub-population t, W_{ut} (m^3/ha), and associated confidence limits, are:

---

[2] The B.C. Forest Service Scaling Manual, 1988, by B.C. Forest Service, Ministry of Forests, Valuation Branch, 1450 Government Street, Victoria, B.C. Canada. V8W 3E7.

$$\hat{W}_{ut} = \sum_{j=1}^{l} H_j T_j \bar{r}_j)/\sum_{j=1}^{l} H_j \ \pm\ t\sqrt{s^2/\sum_{j=1}^{l} n_j} \tag{5}$$

where r_{ij} = ratio of measured volume to estimated volume of residue and waste in accumulation (sample unit) i in stratum j where residue and waste volumes were measured

$\bar{r}_j = \sum_{i=1}^{n_j} r_{ij}/n_j$

s = weighted standard deviation of volume ratios $\{r_{ij}\}$ in all strata

$= \sqrt{\sum_{j=1}^{l} H_j^2 T_j^2 s_j^2/(\sum_{j=1}^{l} H_j)^2}$

s_j = standard deviation of volume ratios $\{r_{ij}\}$ in stratum j

$= \sqrt{\sum_{i=1}^{n_j}[r_{ij} - (\sum_{i=1}^{n_j} r_{ij}/n_j)]^2/(n_j - 1)}$

n_j = number of accumulations (sample units) in stratum j in which the residue and waste volumes were measured

T_j = sum of residue and waste volumes estimated for all 3P units (m^3/ha).

and all other symbols are as previously defined.

AGGREGATED POPULATION

The Aggregated Population can be surveyed by sampling each cut-block as in the Cut-block Population or by using stratified two-stage random sampling.

Two-stage Sampling

Harvested areas in the target population are delineated. Cut-blocks are stratified into block-strata based on cut-block size, timber type, and harvesting system and operator. There is no stratification within cut-blocks. Dispersed and accumulated sub-populations are distinguished in each block-stratum. In the two-stage sampling scheme, the first stage is selection of sample cut-blocks, and the second stage is establishment of fixed-area sample plots within the selected sample cut-blocks as in stratified random sampling.

The number of sample cut-blocks (first-stage units) required, m, is

$$m = \frac{100^2 M t^2 (V + S_w\sqrt{V c_2/c_1})}{M E^2 r^2 + 100^2 t^2 S_b^2} \tag{6}$$

where $V = S_b^2 - (S_w^2/N)$

S_b = standard deviation among cut-block volume means in the population

S_w = standard deviation among plot volume means within cut-blocks in the population

M = total number of cut-blocks in all strata

c_1 = average survey cost per cut-block

N = average total number of plots per cut-block

c_2 = average survey cost per plot in a cut-block

r = estimated mean residue and waste volume in the population (m^2/ha)

E = allowable sampling error defined in equation 1
and t = t-distribution value at the 95% probability level with $m - 1$ degrees of freedom.

It is assumed in equation (5) that the cut-blocks are of equal size; the sampling total cost C is defined by a function of the form: $C = c_1 m + c_2 mn$, where n is the number of sample plots per cut-block; and $V > 0$.

The required number of sample cut-blocks allocated to block-stratum j, m_j, is determined by block-stratum sizes, sampling costs, and variances as

$$m_j = m \frac{H_j S_{bj} / \sqrt{c_{1j}}}{\sum_{j=1}^{q} H_j S_{bj} / \sqrt{c_{1j}}} \tag{7}$$

where c_{1j} = survey cost per cut-block in block-stratum j
S_{bj} = standard deviation among cut-block volume means in block stratum j
H_j = size of block-stratum j (ha)
q = total number of block-strata
and all other terms are as previously defined.

The required number of sample plots per cut-block in each block-stratum, n_j, is determined by stratum sampling costs, and variances as

$$n_j = \sqrt{(c_{1j}/c_{2j}) S_{wj}^2 / [S_{bj}^2 - (S_{wj}^2 / N_j)]} \tag{8}$$

where c_{2j} = survey cost per plot in block-stratum j
S_{wj}= standard deviation among plot volume means in block stratum j
N_j = average total number of plots per cut-block in block-stratum j
and all other symbols are as previously defined.

The selected sample blocks in each block-stratum are subsampled with fixed-area plots as in single-stage sampling. Procedures for measurement and classification of individual pieces of wood are as previously outlined.

Estimates of the mean volume of residue and waste of type u in sub-population t, W_{ut} (m^3/ha), and associated confidence limits, are

$$\hat{W}_{ut} = \sum_{j=1}^{q} H_j \bar{r}_j / \sum_{j=1}^{q} H_j \pm t \sqrt{\sum_{j=1}^{q} (H_j^2 s_j^2) / (\sum_{j=1}^{q} H_j)^2} \tag{9}$$

where $\bar{r}_{jk} = \sum_{i=1}^{n_{jk}} w_{jki} / n_{jk}$
= mean plot residue and waste in cut-block k within block-stratum j (m^3/ha)
$\bar{r}_j = \sum_{k}^{m_j} (N_{jk} \bar{r}_{jk}) / \sum_{k}^{m_j} N_{jk}$
w_{jki}= mean volume of residue and waste in plot i within cut-block k within block-stratum j (m^3/ha)
N_{jk}= possible number of plots located in cut-block k in block-stratum j
m_j = number of sample cut-blocks in block-stratum j
N_j = average possible number of plots per cut-block within block-stratum j
M_j = possible number of cut-blocks in block-stratum j

$f_1 = (M_j - m_j)/M_j$

$f_2 = (N_{jk} - n_{jk})/N_{jk}$

$s_j = \sqrt{f_1(1/(m_j N_j^2))s_{1j}^2 + [1/(m_j M_j N_j^2)] \sum_{k=1}^{m_j} N_{jk}^2 f_2 s_{2jk}^2 / n_{jk})}$

$s_{2jk} = \sqrt{\sum_{i=1}^{n_{jk}} (w_{jki} - \bar{r}_{jk})^2 / (n_{jk} - 1)}$

n_{jk} = number of sample plots located in block k, within block-stratum j

$s_{1j} = \sqrt{\frac{\sum_{k=1}^{m_j} (N_{jk}\bar{r}_{jk})^2 - 2\bar{r}_j \sum_{k=1}^{m_j} (N_{jk}^2 \bar{r}_{jk}) + \bar{r}_j^2 \sum_{k=1}^{m_j} N_{jk}^2}{m_j - 1}}$

t = t-distribution value at the 95% probability level with $\sum_{j=1}^{q} (m_j - 1)$ degrees of freedom

and all other terms are as previously defined.

Note that in equation (6)

S_b is estimated by $\sqrt{\sum_{j=1}^{q} s_{1j}^2 H_j^2 / (\sum_{j=1}^{q} H_j)^2}$

S_w is estimated by $\sqrt{\sum_{j=1}^{q} \sum_{k=1}^{m_j} s_{2jk}^2 h_{jk}^2 / (\sum_{k=1}^{m_j} h_{jk})^2}$

c_1 is estimated by $\sqrt{\sum_{j=1}^{q} c_{1j} H_j / \sum_{j=1}^{q} H_j}$

and c_2 is estimated by $\sqrt{\sum_{j=1}^{q} c_{2j} H_j / \sum_{j=1}^{q} H_j}$

CONCLUDING REMARKS

Quantifying residue and waste from logging operations is difficult and costly. It can, however, be expedited by the fairly straightforward sampling methods presented here. The specifications for these methods had to strike a balance between statistical integrity and the efficiency of field and office procedures.

Several assumptions were made owing to the complexity of the populations being sampled. For example, accumulations, which are highly variable, were treated as homogeneous sub-populations. However, because the volume of residue and waste in accumulations is usually very small relative to the total utilizable volume in a logged area, the impact of any bias resulting from this assumption is minor. Another major simplifying assumption made was that all cut-blocks within a population being sampled based on two-stage sampling are of equal size. The limitations of this assumption are, however, overcome to some extent by the block sample-selection method which uses an area balance check to modify the probability of selecting the next available block as a sample.[3]

The survey methods are presented without detailed derivation or discussion. As well, no comparisons have been made with other methods presented in the literature, such as line-intersect sampling.

Sampling for residue and waste is a large subject, and several aspects such as sampling the accumulations will be studied further. However, the practice of residue surveys in B.C. has over 20 years of experience on a far less rigorous statistical basis than that now developed. The confidence in the estimates provided by the residue surveys will be greatly enhanced by the adoption of these new sampling methods.

ACKNOWLEDGEMENT

We gratefully acknowledge the contribution of the Waste Assessment Committee, especially Chairman Victor S. Dove, of the B.C. Ministry of Forests, and the review comments by J.Thrower and L. Darling of the Research Branch.

---

[3] The block sample-selection method is outlined in the manual referred to in the Introduction.

Critical Height Sampling:
A Workshop On The Current State Of The Technique

Kim Iles
John Bell & Associates
412 Valley Place, Nanaimo, B.C.
Canada V9R 6A6

(corrected copy, corrections in italics)

ABSTRACT

Critical Height Sampling is a method which extends (to three dimensions) "variable plot sampling" or "angle-count sampling" which was developed by Walter Bitterlich in the mid-1940s. It has a few nice properties, which make it interesting from a sampling viewpoint. Specific points of interest are:

- It does not require volume tables.
- It is sensitive to the actual shape of the tree.
- It has useful properties with remeasured permanent sample points, since the estimated volume depends upon tree distance.

Like many systems, the eventual importance of Critical Height Sampling may come from a better solution to the basic intentions of the system, or from some special application of it. For those not familiar with the system, an overview of the ideas involved and the development of the method may be helpful, and is included in this paper. Recent developments and directions for research are covered.

PURPOSE OF THIS PAPER

The purpose of this short paper is to accompany a workshop setting which discussed Critical Height (CH) Sampling. I will not attempt to produce a monograph on the literature and development of CH sampling. Most of the discussion here will be on the use of CH sampling to conduct a simple, one-time inventory. I personally believe that the most promising role of this method is in remeasured permanent sample points, but that has been pretty thoroughly covered in a previous paper (Iles 1987). An earlier paper (Iles 1981) contains a few more details, and the two of them give a reasonable introduction to how CH sampling fits into the permanent sample scene. Since there is a separate workshop on the topic of remeasured point samples during this meeting, I will not fully develop these ideas in this paper.

WHERE IS THE LITERATURE?

A fairly thorough list of Critical Height papers is attached. More than 30 papers have been cited on the subject, from 1962 to 1988, and probably the total is nearer 50 when the complete literature is considered in all languages. I trust that the reader will tolerate the fact that I will quote many of my own works during this short review. I certainly do *not* want to imply that it has been the most important work, but it is the shortest readable series of papers which covers the topic in English for the layman, and I am familiar with the contents and intent. I have tried to reference the person first responsible for any ideas discussed. I apologize if I have unfairly represented anyone's work, and would hope that they would immediately inform me of any corrections.

Some of the most clever work on this topic has been done in Japan, where it was originated, and this has made many of the minor improvements unavailable to non-Japanese readers, myself included.

If anyone knows of earlier developments than those mentioned here, I would appreciate knowing about them so that I can correct any errors in the future. The broader implications of this method are just now being explored, and I will try to confine myself to this general discussion, while giving some of the basic ideas of the classical CH method. Those already familiar with CH sampling may be bored by the following two sections ("Basic Idea of Critical Height Sampling" and "How did it Develop?"), which are inserted for the benefit of the more casual reader of these proceedings.

BASIC IDEA OF CRITICAL HEIGHT SAMPLING

The field work is fairly simple. This method samples the "depth" of wood along random or systematic rays perpendicular to the ground surface. It uses a relascope (usually) to measure this directly, or computes it from an assumed taper of the tree and a known distance. It is a very general idea, just as one might get the volume of snow on an area by measuring the average depth and then multiplying by the horizontal area covered to this average depth. In a more general sense, one could determine the percentage of random lines through a box (containing the forest stand) which actually penetrate wood. If 10% of the rays are passing through wood, then 10% of the box volume would be wood volume. In this particular application, the rays are all vertical. These sampling ideas can be easily proposed. The interesting point is that there are simple ways to actually carry out the job in the field. That makes all the difference.

HOW DID IT DEVELOP?

My part in the development of critical height was an accidental one. Perhaps it is worth recounting, simply because it illustrates one line of thinking which leads to generalizing this method from the two-dimensional method used by Bitterlich into a three-dimensional system which gets volume directly.

In 1974, I was getting a Masters Degree with John Bell at Oregon State University. I was also a teaching assistant, trying hard to explain to a student that Variable Plot Sampling was really a two-stage procedure. The basal area, in two dimensions, was obtained by counting trees. The third dimension, a length, was obtained by measuring some sample trees to get the average height of wood "sitting on" that basal area (called the "average Volume to Basal Area Ratio" or "VBAR"). The two processes, 2-D (area) and 1-D (height) combined to give a 3-D estimate of solid wood volume. "Why, asked my student, couldn't you do it in three dimensions all at once?" Good question ... I hadn't a clue. I did, fortunately, have a lunch hour coming up.

How can you Sample the Solid Volume of a Stand (simply)?

In a straight analogy to the Bitterlich procedure, I visualized the procedure as estimating the percent volume of a transparent "box" which was pierced by random rays. If 30% of the rays penetrating the box were passing through wood, then obviously 30% of the box volume was solid wood.

[Percent wood] * [box volume] = [wood volume]

This is a standard solution to this kind of problem in many fields, particularly the medical field which *has* developed the specialty of "stereology" in recent years. Trees, however, are not easy to orient randomly. Whole stands are even harder. As in so many of these cases, the theory is sterile without the tools and techniques.

How can you do the Measurements?

It was easy to "expand the tree stems" in order to reduce the sampling variability. An angle gauge would expand an entire tree stem, all along its length, just as it expanded the diameter at breast height. This would allow us to sample a stem expanded to a convenient size. We would now be measuring the penetration of rays through that expanded stem rather than through the actual stem. This didn't change the logic, but would make the measurements easier. In addition, the stem was expanded in proportion to its actual taper, so I didn't have to worry about the form of the tree. There are small differences from proportionality between actual and expanded cross-sections when using angle gauges, but these have always proven tolerable in field work.

How, then should the rays penetrate the cube? It seemed obvious that a systematic sample would be preferable, and that a set of rays at 45 degree angles would be the right thing to do. I simply could not think of a way to do it, and only had half a lunch hour left. It was time for a cheap trick.

45 degree rays were difficult. Vertical rays, however, were simple to do - even if they were less efficient statistically. Looking up and down a tree with a Relascope, because of the nature of its ingenious construction, is virtually the same as levitating above and below ground level while viewing the tree horizontally. The number of bars covered by the tree bole, when viewed at an angle from the ground, is exactly the same as if you had risen vertically from your position and were looking horizontally at the tree from a higher vantage point. The magnification may be slightly different, but the relative view is virtually identical.

This was exactly what was needed. When the tree stem was "borderline" (exactly subtending the critical angle), then that was the point where a vertical ray intersected the expanded stem of the tree. Any distance along the stem where the tree was "in" with an angle gauge would represent a portion of the vertical ray which was penetrating the expanded tree stem. In many cases, if we were lucky, the tree stem would be constantly decreasing, and there would only be one "penetrated distance". If not, we could add them together. In addition, there may be several trees involved, since the "expanded stems" can overlap each other. We would therefore add up all the segments for all trees. Figure 1 illustrates the geometry.

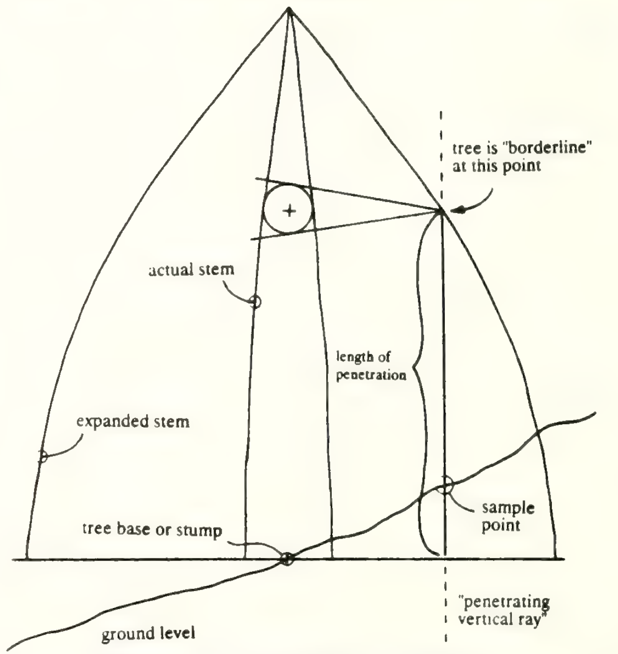


Figure 1. Basic geometry of system.

What is the Stand Volume?

Since the trees were "expanded" by the angle gauge (say 1000 times their original volume), we are obviously getting the average depth of "expanded tree stems". Therefore, the "average depth" would have to be reduced by the amount of the expansion (1000 times). You then had the "average depth of solid wood" on the land base. If you wanted the volume on a hectare, then you just multiplied by that average depth by the area of one hectare.

average depth * area = estimated volume

It just so happens that there is already a factor to do these two adjustments. The Basal Area Factor (BAF) of a prism does both these tasks at once, since they are both constants. It reduces for the "expansion factor" and multiplies by an area with particular units (square feet per acre, or square meters per hectare for instance.)

As an example, we only have to add up the sum of the penetrated segments, then multiply by the BAF of the angle gauge used.

total length * BAF = volume/acre (or hectare)

This gives an estimate at each sample point. No other tree measurements are necessary, just the length of penetration. If you want log grades, just record them by segment lengths. Decay is handled the same way. Unmerchantable portions of the tree are simply ignored (unless you later want an estimate of their volume, in which case you just don't use them in the net volume computations).

What is the Sampling Error?

Since you have an estimate of area volume at each sample point, the variance estimator is obvious. It is proportional to the sum of the penetration segments, just as it would be based on the sum of VBARS per point in ordinary variable plot sampling. Both of these measurements are implying the solid wood volume above a particular sample point.

In essence, the average penetrated length per tree is the same concept as average VBAR for the tree, but it is based on a sample of the actual stem, rather than computed from extra measurements and assumed volume equations. Another difference is that the basal area is from the stump with this new system (rather than at DBH as it normally is with variable plot sampling).

So much for the lunch hour ... I now had a system which extended variable plot sampling to the third dimension, used actual tree shape, required only one measurement to unbiasedly get tree volume (bark included), and could use existing instruments. I called it "penetration sampling", and wrote up a small report for one of my classes. The general reaction was that: (a) it couldn't be that simple, and was therefore in error; or, (b) if it was that simple, it had surely been invented already.

Was the System New?

It was hard to believe so. It was just too simple. Still, it took about a month of hard work to find even a mention of it, and that was to the credit of "Forestry Abstracts" which had a very brief description of a Japanese paper translated by the Canadian Department of Fisheries and Forestry. It contained the cryptic phrase "the sum of critical heights". Since "critical angle" is a common term in variable plot sampling, it was quite possible that this might refer to the heights to where the stem just subtended that critical angle. This proved to be the case. I obtained a copy of the original Japanese article, and it contained a figure virtually identical to the ones I had been drawing. Later examination of the translation, and other articles, proved that the systems were identical in all important features.

The author, Dr. Kitamura, had begun to publish articles on the method more than a decade previously, but the system was virtually unknown in North America. Since he has clear priority for this method, I have attempted to use his terms whenever possible. He calls the penetrated length the "critical height" for a tree (I think the extension to adding segments is obviously implied). The German term is "Deckpunkthohen". Dr. Kitamura has written some articles in German, but most of them are in Japanese. For those interested, I have some partial translations available for some of these papers.

DOES CH SAMPLING WORK WELL?

Visibility

My first field tests of the system were at the University of British Columbia, as one part of a PhD thesis. Visibility inside the crown of Northwest conifers did not cause many problems, and those could be avoided by several methods. Steep sighting angles were also a minor problem, but this could be solved in several ways also.

Specifying the "Critical Point"

There were two main problems with the system. First, the trees were far more "bumpy" than expected. It was not quite trivial to locate the critical point on the tree. We are accustomed to thinking of trees as tapering monotonically, but this in not the case. Remember that with a 40 BAF prism each foot of critical height represents 40 cubic feet of wood. Even magnification does not cause this problem to go away. One solution to this problem is to use an indirect method for establishing critical height. By assuming a tree taper for the tree (after taking enough measurements to establish tree volume accurately) you can calculate a "pseudo-tree" critical height. This insures that the tree taper is monotonic. This approach was suggested by Kitamura (1968 and possibly sooner), and explored by Bitterlich (1976) and Iles (1979b). You have two main choices: a completely arbitrary tree shape using a volume table (which introduces various possible biases); or, you can smooth out the tree from actual measurements (which mainly involves interpolation accuracy).

There is also one additional solution. You could use a counting system near the critical point. Often I encountered small segments and "gray zones" where I had trouble deciding upon the exact critical point, but felt fairly comfortable counting particular points "in" or "out". A practical solution might therefore be to make discrete decisions near the critical point by testing a series of points. The percent scale of a relascope can easily choose equidistant points to make this easy. Both Tom Beers and Mike Bonner had informally suggested this kind of scheme to me in about 1975, but in a slightly different context. It nicely solves the problem of small "critical segments" which occur from the natural swelling of the bole - particularly near branch whorls. It was the only practical solution I could suggest, but I suspect that smoothing with an interpolation method, such as Bitterlich suggests, would be better. Ueno has authored a paper of a similar nature, but taking random counts along the entire bole.

These various schemes sometimes lost the advantage of an unbiased volume determination, and complicated the measurements required. They retained the advantage of using the distance to the tree when estimating stand volume. The hope was that the natural tendency of trees to space themselves and to react to nearby competition would result in less variable volume estimates.

Comparison to the Usual Method

Unfortunately, the second problem was worse. A simulation, using an actual stem map, indicated CH sampling to be just as variable as ordinary variable plot sampling. The results were better than if the trees were randomly placed, but there was clearly no great advantage in variability to be gained because of inter-tree spacing. Details are available in Iles (1979a, 1979b). This lead to a search for other uses for the system.

USE WITH PERMANENT VARIABLE PLOTS FOR GROWTH ESTIMATION

One of the first uses to be explored was with permanent sample points. With the usual techniques up to that time, new trees which appeared on the sample point at a second measurement were either ignored, or they caused a large and disruptive change to the growth estimate. The choice was one of "incompatibility" *vs* large variation. While I personally believe that the question of "incompatibility" is largely a red herring, CH techniques offer a nice alternative. In addition, the growth estimates are taken directly from the tree, and are sensitive to tree form. This essentially trades a measurement precision problem (with CH) for concern about a volume table bias (with the usual systems). The change in the sum of critical heights from one period to another will offer a direct estimate of stand growth during the period. Figure 2 below shows the basic idea. These points are detailed in the papers listed earlier. A short time after this proposal was made, (Iles 1979a, 1979b) a similar suggestion was independently derived by Flewelling (1981).

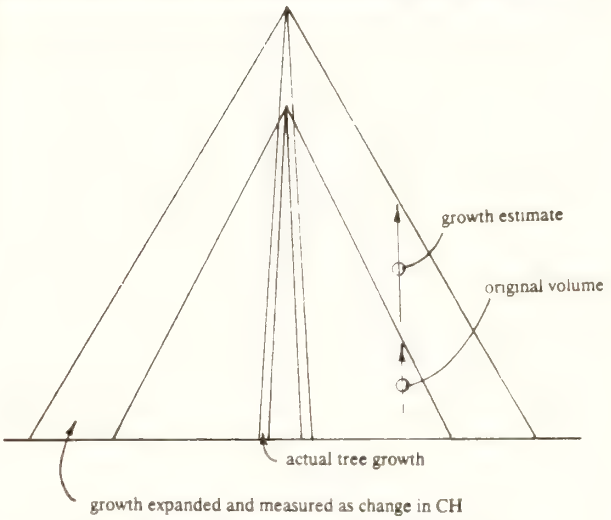


Figure 2. Critical Height for growth estimation.

WHAT WERE THE ANCESTORS OF THE METHOD?

It seems quite remarkable that the system was not discovered earlier. Bitterlich actually came very close to it with a paper in 1956, where he took basal area sweeps at several heights in the stand and calculated the volume from these basal areas. CH is just a continuous form of this discrete systematic sample. Ueno (1979) used a random sampling method to count trees at various random heights from the ground. This seems to be a step back from Bitterlich's systematic sample, unless there is a mechanism for choosing the random points between the ground and some fixed height with very little effort. I know of no such system but, in fairness to Ueno, I must point out that I have not read his original paper.

RECENT DEVELOPMENTS, NEW SUGGESTIONS

Form Factor

It is quite simple to measure, or sample (systematically or randomly), if you are considering only relative tree height. The relascope, using the "percent scale" (a tangent of the angle), can easily pick points along the bole that are equal distances apart. We do not need to know the distance to the tree. The first even 5% or 10% mark will be virtually a random start. This offers a quick mechanism for determining the form factor of the tree. Recording the total "% height of the tree", and then counting how many random or systematic points along the stem are "in" with the angle gauge, will indicate the percent cylindrical form factor. The critical height is just the continuous analog of this process.

Selecting trees within a cluster would be unbiased, and also less variable than doing this by individual trees. Selection by use of the prism would give an automatic weighting by basal area. I believe that an additional weighting by estimated height would also be desirable. I do not know of anyone who has previously suggested this procedure, but it would not be surprising it has been done.

Sub-sampling

It is easy to see the application of sub-sampling to the system. Since the average critical height per tree is an analog to the average VBAR in ordinary variable plot sampling, the methods commonly used to adjust volumes by the use of "count-only" plots can be used with virtually no changes. Simple tree counts, with the same prism at extra points, can thereby increase the precision of the sample. From a translation of the title page of one of Kitamura's papers I know that he discusses double sampling, and presume that this idea was first discussed by him, but I know of no direct citation.

Possible Connections to "Importance Sampling"

We could also estimate individual tree volumes with a critical height procedure by averaging the critical height of the tree from a variety of random points around the tree. This has been suggested from the beginning of the system. In order to considerably sharpen the estimate, we could make use of a ratio estimator. By first using an estimate of the critical height from any reasonable taper function, then using the actual critical height to calibrate the estimate. It seems to me that this is the basic suggestion of "Importance Sampling", but done in a vertical, rather that horizontal fashion. The workshop on "Importance Sampling", which is being held at this conference, can more competently address this issue than I can.

Modifying Functions

VanDeusen (1987) has suggested a modification to the CH procedure which is worth reviewing. Essentially, the suggestion is to add a "modifying number" to the original critical height. The modifier has an average of zero, so the unbiased nature of the direct critical height volume estimate is unaffected, but the estimate itself may be less variable from the edge of the plot to the center. Essentially it is a way to change the shape of the tree, but retain the correct volume. Whether his particular use of this method is an advantage or not may be debatable. The reason that I am partial to this suggestion is that it leads in a direction that I think is constructive, not only for CH sampling, but for sampling in general.

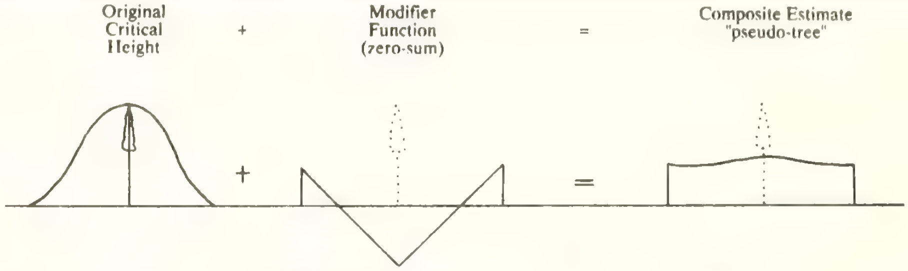

Original Critical Height
+
Modifier Function (zero-sum)
=
Composite Estimate "pseudo-tree"
+
=

Figure 3. Using modifier function to create "pseudo-tree" shape.

GENERALIZATION OF THE PROBLEM

In 1983, I gave a talk at the Western Mensurationist's Meeting with the title "Shiftable Plot Borders, Pseudo-trees and the CFI Concept. Why not take stem distribution into account?" As it concerned Variable Plot sampling systems of various types, the basic message was this:

> For years now, we have used estimators (once a tree was selected) which were essentially "flat" or disc-shaped. A tree represented the same volume per acre if were 5 or 50 feet away. Still, we believe that trees naturally space themselves out and react to the space made available to them. Often we have distances to these trees, and even inter-tree distances. Why not make the estimators vary across their surface area? We can use the expanded natural shape, or change it, by various analytic methods into "shaped estimators" or "pseudo-trees" which are not quite the shapes of the original trees. If we are farther from a tree, it would imply less volume per acre at the point where we are sampling than an identical tree which was closer to us. Not only does this appeal logically, but it may have statistical advantages as well.

One could also think of the different value of the estimator at different distances from the tree as arising from a "weighting procedure" applied to the average value.

What are Examples of This?

The indirect methods of CH sampling are an example of changing the actual tree shape into a form more smooth and predictable. We may find that other forms result in less variable stand volume estimators than the tree forms themselves. I believe that we have been unimaginative in using distance-dependent estimators. As measurement technology changes, distance to trees becomes easier to record. With computers, the interpolation process becomes reasonable. The method of "Regionalized Variables" or "Kriging" takes a similar approach in that it creates a "smoothed" surface rather than a "stepped" one, across the sampled area. Some methods of getting the variance of systematic samples take an essentially similar approach.

Although the VanDeusen suggestion makes the tree estimator "flatter" over its selection area, the important point is that it offers yet another way of controlling the shape of as estimator, and is therefore a useful tool. One can think of many ways to use combined estimators in a similar vein, all of which *seem* unexplored. The benefits might accrue not only to fixed plots, but sampling systems quite different that we usually use in forestry. One of the interesting points about Critical Height sampling is that it is moving us in that direction.

Further Work

The system is badly in need of field testing. There are a few prime applications:

1) Trees without volume tables (or with taper that is very unreliable - hardwoods for instance, which also have good visibility in the winter).

2) Remeasured sample points (especially where distances to trees are already available).

It may also be possible that CH can be desirable as one of the estimators in a mixed estimation system. The problem of bark volume must be solved to the satisfaction of the inventory units considering this use, but I would assume that this is no more serious a problem than they currently are facing with bark thickness. Yang (1983), has already done some of the work in extending the basic principles of this work to horizontal line sampling. This particular adaptation is almost identical in nature to "line intersect" sampling.

REFERENCES

Beers, T.W. and C.I. Miller. 1964. Point sampling: research results, theory and application. Purdue Univ. Agric. Exp. Stn. Res. Bull. 786, 56 p. Lafayette, IN.

Bitterlich, W. 1975. Volumsstichprobe aus indirekt bestimmten Deckpunkthohen. Allgemeine Forstzeitung 86(4), 113-5.

Bitterlich, W. 1976. Volume sampling using indirectly estimated critical heights. Commonwealth Forest Review, Vol. 55, No. 4, pp. 319-330.

Bitterlich, W. 1983. Eine fliessend mitwachsende Dauerstichprobe im Wald. Allgemeine Forstzeitung 94 (2):33-4.

Flewelling, J.W. 1981. Compatible estimates of basal area and basal area growth from remeasured point samples. Forest Sci. 27:191-203.

Flewelling, J.W. and C.E. Thomas. 1984. An improved estimator for merchantable basal area growth based on point samples. Forest Sci. Vol. 30, No. 3, pp. 813-821.

Grosenbaugh, L.R. 1958. Point-Sampling and Line-Sampling: Probability Theory, Geometric Implications, Synthesis. USDA Forest Service, Southern Forest Experiment Station Occasional Paper 160. 34 pp.

Grosenbaugh, L.R. 1959. Should Continuity Dominate Forest Inventories? Paper presented at Short Course on Continuous Inventory Control in Forest Management. May 4-8, 1959. Athens, Georgia.

Iles, K. 1974. Penetration sampling--an extension of the Bitterlich system to the third dimension. Oregon State University, Corvallis. Unpublished manuscript.

Iles, K. 1979a. Some techniques to generalize the use of variable plot and line intersect sampling. In: Forest resource inventories workshop proceedings (W.E. Frayer, ed) v I, p. 270-278. Colo. State University, Fort Collins. 2 v, 1037 p.

Iles, K. 1979b. Systems for the selection of truly random samples from tree populations and the extension of variable plot sampling to the third dimension. PhD diss, University of British Columbia, Canada. 144 p. (Diss Abstr. 40/04B.)

Iles, K. 1981. "Permanent 'Variable' Plots for Forest Growth". Western Forest Mensurationists. Sun Valley, Idaho.

Iles, K. and Tom W. Beers. 1983. "Growth Information from Variable Plot Sampling." Int'l Conf. Renewable Resource Inventories for Monitoring Changes and Trends, Corvallis, OR (J.F. Bell and T. Atterbury, eds.).

Iles, K. 1987. Growth on permanent angle-count plots, historical procedures up through critical height sampling. IUFRO Forest Growth Modelling and Prediction Conference. Minneapolis, MN, U.S.A.

Kitamura, M. 19??. Estimation of the stand volume through the sum of critical heights, IUFRO, section 25. (In German, translation available--Iles.)

Kitamura, M. 1962. On an estimate of the volume of trees in a stand by the sum of critical heights. 73 Kai Nichi Rin Ko, p. 64-67. (In Japanese, translation available--Iles.)

Kitamura, M. 1964. Theoretical Studies on the Estimation of Stand Volume Through the Sum of Deckpunkthohen. Bulletin of the Yamagata University (Agricultural Science), Vol. 4, No. 4, pp. 365-403. (In Japanese, table of contents translated--Iles, also short English summary.)

Kitamura, M. 1965. Estimate of the sum of critical heights by Spiegel Relascope, Collection of the 76th Lectures held for the Japanese Forestry Society. (In Japanese, English translation available--Iles.)

Kitamura, M. 1968. A simple method of estimating the volume of trees in a stand by the method of the sum of critical heights. Transl. Dep. Fish. For. Can., 1969 OOFF-3, 19 p. Transl. from J. Jap. For Soc. 1968, 50 (11):331-335.

Kitamura, M. 1976. Zur Vereinfachung der Methode durch die Deckpunkthohensumme. Rep. 28th meeting Tohoku Bra. Jap. Forestry Soc. (In Japanese, translation available--Iles.)

Kitamura, M. 1977. Eine Methode zue Bestandesmassenermittlung mit dem vertikalen kritischen Winkel. Journal of the Japanese Forestry Society, Vol. 59, No. 3, pp. 101-103. (In Japanese, translation available--Iles.)

Kitamura, M. 1978a. Moglichkeiten der Bestandesmassenermittlung mit dem horizontalen oder verticalen kritischen winkel. IUFRO proceedings, Nancy, France. (In German, English summary.)

Kitamura, M. 1978b. Untersuchungen uber die Bestandesmassenermittlung durch die Deckpunkthohensumme, Mitteilungen vom Institut fur Forsteinrichtung, Universitat Yamagata. (In German.)

Lynch, Thomas B. 1986. Critical Height Sampling and Volumes by Cylindrical Shells. Forest Sci., Vol. 32, No. 3, pp. 829-834.

Martin, G.L. 1982. A method for estimating ingrowth on permanent horizontal sample points. Forest Sci. Vol. 28, No. 1, pp. 110-114.

McTague, J.P. and R.L. Bailey. 1985. Critical height sampling for stand volume estimation. Forest Sci. Vol. 31, No. 4, pp. 899-911.

Rubinstein, R.Y. and G. Samorodnitsky. 1985. Variance reduction by the use of common and antithetic variates. J. Statist. Comput. Simul. 22:161-180.

Ueno, Y. 1979. A New Method of Sampling Stand Volume. I. [in japanese] Journal of Japanese Forestry 61 (9).

Van Deusen, Paul C., Tommy R. Dell, and Charles E. Thomas. 1986a. Volume Growth Estimation from Permanent Horizontal Points. Forest Sci., Vol 32, No. 2, pp. 415-422.

Van Deusen, Paul C. and Walter J. Meerschaert. 1986b. On critical-height sampling, Can. J. For. Res., Vol. 16, pp. 1310-1313.

Van Deusen, Paul C. and Thomas B. Lynch. 1987. Efficient Unbiased Tree-Volume Estimation. Forest Sci., Vol. 33, No. 2, pp. 583-590.

Y.C. Yang and C.H. Yang. 1983. Expectations and variances of basal area estimates using remeasured horizontal line sampling. Forest Sci., Vol. 33, No. 1.

Sampling and Modelling Visual Component Dynamics of Forested Areas

Victor A. Rudis
Research Forester, Forest Inventory and Analysis Unit
USDA Forest Service, Southern Forest Experiment Station
201 Lincoln Green, Starkville, MS 39759-0906

ABSTRACT

A scaling device and sample design have been employed to assess vegetative screening of forested stands as part of an extensive forest inventory. Referenced in a poster presentation are results from East Texas pine and oak-pine stands and Alabama forested areas. Refinements for optimizing measures to distinguish differences in scenic beauty, disturbances, and stand structure are discussed.

SUMMARY

Estimating verifiable and repeatable judgment of vegetative screening for studies of forest aesthetics, hiding cover for wildlife, and as a measure of forest disturbance has been difficult, particularly in field surveys of extensive areas. A scaling device and sample design have been developed to estimate within-stand visual components in forested areas: visual penetration, tree bole screening, and foliage, twig, and small stem screening. The scaling device, a "screenometer," is a clear plastic plate (7.5 inches by 2 inches), scored with divisions, and held 14 inches from the eye. Visual components are estimated at fixed distances from the observer.

Methods are discussed in Rudis (1985) and Rudis and others (1988). Data are presented in Rudis and others (in preparation). Associated with this study, but published elsewhere, are reports on a dispersed recreation inventory of Alabama's forests (Rudis 1983), recreation use and recreation opportunity preferences of Alabama forest users (Rudis 1987), and psychological utility of visual penetration measures in scenic beauty models (Ruddell and others 1989).

Sampling at a 50-foot radius from an observer position with a "screenometer" provides a minimum distance to distinguish forested plots with and without recent disturbances and to maximize correlations with scenic beauty valuation in East Texas pine and oak-pine stands.

Data from a broad-scale survey of 3,700 forested plots in Alabama suggest a conceptual model of visual components over time as follows: a rapid decline in visual penetration as vegetation approaches viewing height, followed by a gradual rise in visual penetration with increasing stand age and average stand diameter. As stands mature, expanding tree canopies and periodic disturbances reduce foliage, twig, and small stem screening. Tree bole screening is negligible to age 8 in even-aged stands -- or to an average stand diameter of 2 inches at breast height among all-aged stands. Tree bole screening increases with stand age and larger average stand diameter. Stands with recent cutting, livestock grazing, and recent burning have more visual penetration than stands without such disturbances.

Ocular estimation of vegetative screening is used commonly to estimate aesthetic impacts of forest management practices, to classify recreation opportunities, and to assess wildlife hiding cover and understory biomass. With the aid of a "screenometer," vegetative screening can be objectively inventoried and quantified. Differences at 50 feet from the observer are more closely associated with differences in scenic beauty and stand disturbances than shorter distances, and are logically related to differences in stand structure.

The broad-scale survey of visual components in Alabama forests provides an important conceptual model for regional simulation. Variability among stands, however, is sufficiently diverse to preclude model prediction of visual vegetative screening based on stand age, average stand diameter, and stand disturbances. Subdivision of screenometer sectors, inventory of more detailed components, and better knowledge of stand establishment, timing and duration of disturbances, and environmental conditions, should improve prediction of vegetative screening from other survey measures in sampled areas.

REFERENCES

Ruddell, E.J.; Gramann, J.H.; Rudis, V.A.; Westphal, J.M. 1989. The psychological utility of visual penetration in forest scenic beauty models. Environment and Behavior 21(4): 393-412.

Rudis, V.A. 1983. Dispersed recreation inventory on commercial timberland. IN: Bell, J.F.; Atterbury, T., eds. Renewable resource inventories for monitoring changes and trends. 1983 August 15-19; Corvallis, OR. Corvallis, OR: Oregon State University, College of Forestry: 214-218.

Rudis, V.A. 1985. Screenometer: a device for sampling vegetative screening in forested areas. Canadian Journal of Forest Research 15: 996-999.

Rudis, V.A. 1987. Recreational use of forested areas by Alabama residents. Research Paper SO-237. New Orleans, LA: USDA-Forest Service, Southern Forest Experiment Station. 37 p.

Rudis, V.A.; Gramann, J. H.; Ruddell, E.J.; Westphal, J.M. 1988. Forest inventory and management-based visual preference models of southern pine stands. Forest Science 34(4): 846-863.

Rudis, V.A.; Gramann, J.H.; Ruddell, E.J. (in preparation). Visual component sampling to monitor within-stand scenic beauty and stand dynamics from forest inventories. Submitted to Forest Science.

Model-Based Sampling

John C. Rennie
Associate Professor
Department of Forestry, Wildlife and Fisheries
The University of Tennessee
Knoxville, Tennessee 37901-1071

This workshop module was designed to introduce forest biometricians and inventory specialists to model-based sampling and some of its uses in forest resource assessment.

The goal of a sampling scheme is to effectively estimate population parameters. Lacking information on the population, simple random sample with simple expansion estimates is used. With auxiliary data, two approaches are possible: design-based sampling and model-based sampling. Design-based sampling strategies traditionally use the auxiliary data in techniques such as stratification, PPS or ratio and regression estimation. Although a model is tacitly assumed, correctness of the inference does not depend upon it when a reasonably large sample is taken.

Model-based strategies use a model that describes the relation between the auxiliary variable and the variable of interest. Optimal selection and estimation schemes are then developed from this model. More care must be used in applying model-based sampling. If the model is not appropriate, selection and estimation schemes can be far from optimal.

In this workshop I have shown how to start with the combined-variable model for individual tree volume (V_i in equation 1) and derive a model for stand volume (V in equation 2). The number of trees in the population, N, and mean diameter squared height ($\overline{D^2H}$) are estimated from plot, line or point samples. The regression coefficients, a and b, are estimated from trees purposefully selected from the population.

$$V_i = a + b\,(D^2H) \qquad (1)$$

$$V = N\,[\,a + b\,(\overline{D^2H})\,] \qquad (2)$$

Point-model based sampling for timber inventory was discussed in detail. Planning the inventory included defining goals, determining sample size, preparing forms and assembling equipment. Executing the inventory included locating sampling points, selecting sample trees to estimate the number of trees in the population and mean D^2H, and selecting second-stage trees for volume determination. Processing point-model based inventory data included review of the output from a computer program, PTMODEL, to process point-model base inventory data.

Hans T. Schreuder of the U.S. Forest Service's Multiresource Inventory Project, Fort Collins, Colorado, discussed the development of model-based sampling in forest resource assessment. He presented the results from a simulation study estimating the sum of D^2H. Twenty sampling points were randomly located at which trees were selected proportional to D^2; 1 or 2 trees were selected at each sampling point. Three estimators were compared: one with probability inverse weighting, one using weighted regression and one with probability inverse weighting and using weighted regression. Unweighted and weighted jackknife standard errors were calculated and expressed as a percentage of the iterated standard error for each estimator. The jackknife standard errors were close to the iterated standard error for all three estimators.

Charles E. Thomas of the U.S. Forest Service's Institute for Quantitative Studies, New Orleans, Louisiana, reviewed three papers that compare probability sampling and model-dependent sampling (Royall and Cumberland 1981a and 1981b; Hansen et al., 1983). He then discussed a paper by Van Deusen (1987) in which the unadjusted and the adjusted 3P sampling estimators are examined as examples of a design-based estimator and a model-based estimator, respectively.

REFERENCES

Hansen, M.H., Madow, W.G. and Tepping, B.J., 1983: An evaluation of model-dependent and probability-sampling inferences in sample surveys. Journal of the American Statistical Association, vol. 78, no. 384, pp. 776-807.

Royall, R.M. and Cumberland, W.G., 1981a: An empirical study of the ratio estimator and estimators of its variance. Journal of the American Statistical Association, vol. 76, no. 373, pp. 66-88.

Royall, R.M. and Cumberland, W.G., 1981b: The finite-population linear regression estimator and estimators of its variance - an empirical study. Journal of the American Statistical Association, vol. 76, no. 376, pp. 924-930.

Van Deusen, P.C., 1987: 3-P sampling and design versus model-based estimates. Canadian Journal of Forest Research, vol. 17, no. 2, pp. 115-117.

Importance Sampling for Volume with a Portable Computer

Harry T. Valentine
Research Forester
Northeastern Forest Experiment Station, USDA Forest Service
Hamden, Connecticut 06514

Timothy G. Gregoire
Assistant Professor of Forest Biometrics
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

George M. Furnival
J. P. Weyerhaeuser, Jr., Professor of Forest Management
School of Forestry and Environmental Studies, Yale University,
New Haven, CT 06520

ABSTRACT

A two-stage procedure is described that uses list sampling with probability proportional to size and importance sampling to furnish an estimate of the total bole volume of the trees in a forest survey plot. The procedure can be programmed to run in a battery-powered, portable computer. The computer gives instructions to the field crew, performs all of the needed calculations, and furnishes the estimate of volume and its standard error. The procedure is suggested for use in surveys where accurate, unbiased volume equations are unavailable for the species that are likely to be encountered.

INTRODUCTION

We have formulated and computerized a two-stage procedure that furnishes estimates of total bole volume and its standard error for the trees in a forest survey plot. The first stage uses list sampling with probability proportional to size (LS) and the second stage uses importance sampling (IS), a continuous analogue of LS (see: Valentine *et al.* 1984; Furnival *et al.* 1986; Gregoire *et al.* 1986, 1987). Implementation of the procedure requires a durable, lightweight, battery-powered, personal computer. The computer is programmed to give instructions to the survey crew at the plot, perform all computations for the two-stage sampling, and furnish the final volume and error estimates. In this paper, the tasks and sampling formulae involved in the procedure are outlined in sufficient detail to enable a scientific computer programmer to create a program for implementing the procedure. We also offer our own computer program for this purpose.[1] We have attempted to keep our program as simple and as short as possible; it is not intended to do all things for all surveys. It is a kernel program that is intended to be modified or elaborated by a user if need be.

Theory and applications of LS, our first-stage sampling technique, can be found in most sampling texts. The present application is identical to one given by Valentine *et al.* (1987) except that volume of the survey plot is the target parameter instead of dry matter. We use LS because it is an efficient technique and, in the absence of measurement errors, it furnishes an unbiased estimate of the volume of the survey plot and its standard error. Likewise, IS furnishes unbiased volume estimates for the trees selected in the first-stage sample. Our application of IS was developed by Gregoire *et al.* (1987). The joint application of LS and IS requires (1) the use of a cross-sectional area taper equation or, preferably, a volume equation and a taper equation, and (2) a device or instrument to measure upper stem diameters.

---

[1] This computer program is coded Microsoft (MS) Compiled BASIC. A copy of the source code can be obtained via BITNET from T.Gregoire:TGG@VTVM1, or via the Forest Service DG network from H.Valentine:S24L07A.

Our computerized sampling procedure for volume estimation can be used in any survey where trees have clearly defined boles, but it is especially appropriate when direct application of the volume equation(s) at the disposal of the survey team are known or suspected to furnish biased estimates (i.e., estimates of tree volumes that are, on average, too large or too small). Bias may arise if an equation is fitted for one particular population of trees and is used for another. For example, volume equations developed for a large region may give seriously biased results when applied to a specific forest within the region. Or, more obviously, equations developed for one forest (species) may give biased results when applied to a different forest (species). In our procedure, volume and taper equations are used as auxiliary information to increase sampling efficiency; the unbiasedness of our estimators is not affected by the origin or accuracy of the equations.

FIRST-STAGE SAMPLE

A volume or taper equation is used to predict the bole volume of each of the N trees on a survey plot. Therefore, the survey crew must measure the independent variable(s) of the volume (or taper) equation on each tree, e.g., dbh and perhaps height. In our computer program we use a volume equation that requires only dbh. The user has the option to enter an identification code or species code or both into the computer together with the dbh measurement for each of the N trees. Trees (in temporary plots) without identification codes are assigned identification numbers by the computer. These numbers can be written on masking tape and attached to the trees when the variables for the volume equation are measured. Upon completion of data entry, the computer selects a user-specified number (n) of the trees at random (with replacement) with probability proportional to predicted volume, and returns their identification codes or tree numbers. These n trees constitute the first-stage sample.

To select the sample, the computer accumulates the predicted volumes of trees 1 through j (C_j) for $j = 1, 2, \ldots, N$; i.e.,

$$C_j = \sum_{i=1}^{j} y_i, \qquad j = 1, 2, \ldots, N \tag{1}$$

where y_i is the predicted volume of tree i ($i = 1, 2, \ldots, N$). $C_0 = 0$ also is defined. Next, n pseudo-random numbers are generated from a [0,1] uniform distribution. Tree i is selected as the jth ($j = 1, 2, \ldots, n$) tree in the first-stage sample if:

$$C_{i-1} < u_j C_N < C_i \qquad i = 1, 2, \ldots, N \tag{2}$$

where u_j is the jth of the n pseudo-random numbers.

We have used the simplest sort of volume equation in our computer program to calculate the y_i, *viz.*,

$$y_i = aD_i^b \tag{3}$$

where D_i is diameter breast height, a is a scaling constant, and b is an allometric constant. Eqn (3) is convenient because it does not require a height measurement. The constant a can be assigned a value of 1.0, and b can be assigned a value of 2.5 (White 1981) if better information is not available. With such assignments, the y_i will be positively correlated with the true volumes of the trees, but due to the arbitrary scaling (i.e., $a = 1$), they will not be accurate predictions. However, the success of LS depends only on the degree of correlation, and not at all on the scaling.

Many and perhaps most volume equations use height as an independent variable. Volumes also can be predicted by integrating taper equations with respect to height. If an accurate volume or taper equation which uses height is available for a species of interest, then it could be coded into our computer program in place of eqn (3). Visually estimated heights should suffice for the calculation of the y_i.

SECOND STAGE-SAMPLE

Importance sampling is used to estimate the true volume of each of the n trees in the first-stage sample. The computer directs the survey crew to measure either the total height (H) and two intermediate heights (L and U) of each sample tree (by identification code) or distances and slopes for their calculation. The intermediate heights, respectively, are the lower and upper heights of the section of the bole that contains the volume of interest. If $L = 0$ and $U = H$, then the section of interest is the entire bole. Alternatively, L might be stump height, and U, merchantable height. The heights or distance and slope measurements are entered into the computer, which returns an importance sample consisting of a user-specified number (m) of cross sections of the bole at heights between L and U. The diameters of the the bole at these heights must be measured and entered, after which the computer returns the estimate of the bole volume for the tree. This procedure is repeated for each of the n trees in the first-stage sample.

The procedure for selecting an importance sample of m cross sections of the bole at heights between L and U requires an integrable taper equation that predicts the cross-sectional area of the bole, $S(h)$, at any height h ($L < h < U$).

Let θ_k ($k = 1, 2, \ldots, m$) denote the height of the kth cross section in the importance sample. This height is determined by solving the following equation for θ_k:

$$\int_L^{\theta_k} S(h)\mathrm{d}h = u_k \int_L^U S(h)\mathrm{d}h \qquad k = 1, 2, \ldots, m \tag{4}$$

where u_k is a uniform [0,1] pseudo-random number. In our computer program, we use the following taper equation for $S(h)$:

$$S(h) = B\frac{H-h}{H-H_b} \tag{5}$$

where B is basal area (i.e., cross-sectional area at breast height) and H_b is breast height. Substituting eqn (5) into eqn (4), integrating, and solving for θ_k gives:

$$\theta_k = H - \sqrt{(H-L)^2 - u_k[2H(U-L) + (L^2 - U^2)]} \tag{6}$$

Thus, for this particular taper function, the value of θ_k depends only on the heights L, U, and H that are measured by the survey crew, and the random number, u_k, that is generated by the computer.

Although eqn (5) integrates in closed form and furnishes the relatively simple eqn (6) for θ_k, we do not use this latter equation in the computer program. Anticipating that many users may want to substitute their own taper equation for eqn (5), we have developed a simple numerical procedure that furnishes θ_k for any taper equation $S(h)$, no matter how complicated.

In our numerical integration procedure, we assume that the bole of a tree between L and U can be modeled by a stack of M wafers, each with thickness Δh, (i.e., $M = (U-L)/\Delta h$, where Δh equals 1 cm). The cross-sectional area of each wafer is calculated with the taper equation $S(h)$, depending on its midpoint height within the bole. The cross-sectional area of the ith wafer (S_i) is:

$$S_i = S[L + (i - 1/2)\Delta h] \qquad i = 1, 2, \ldots, M \tag{7}$$

A new segmented taper equation, $S^*(h)$, that describes the profile of the wafer model of the bole of the tree is defined from the S_i as follows:

$$S^*(h) = S_i \qquad \begin{array}{l} L + (i-1)\Delta h < h < L + i\Delta h \\ \text{for } i = 1, 2, \ldots, M \end{array} \tag{8}$$

Thus, S_i equals the predicted cross-sectional area at all points within the bole between heights $L + (i-1)\Delta h$ and $L + i\Delta h$, i.e., within the ith wafer.

To find θ_k which satisfies eqn (4) for $S^*(h)$, we first calculate:

$$T_i = \sum_{j=1}^{i} S_j \qquad i = 1, 2, \ldots, M \tag{9}$$

and define $T_0 = 0$. The height θ_k occurs in the ith wafer if:

$$T_{i-1} < u_k T_M < T_i \qquad i = 1, 2, \ldots, M \tag{10}$$

Given the index i of the wafer in which it occurs, θ_k is calculated as:

$$\theta_k = L + \left(i - 1 + \frac{u_k T_M - T_i}{S_i}\right)\Delta h \tag{11}$$

The predicted cross-sectional area of the bole at θ_k within the ith wafer is:

$$S^*(\theta_k) = S_i \tag{12}$$

VOLUME AND VARIANCE CALCULATIONS

An estimate of the volume of each tree in the first-stage sample is calculated by the computer from the diameters measured at heights θ_k ($k = 1, 2, \ldots, m$). Let $d(\theta_k)$ denote diameter at height θ_k on the bole of tree i. Assuming the bole is round, the actual cross-sectional area at θ_k is:

$$A(\theta_k) = \frac{\pi}{4} d(\theta_k)^2 \tag{13}$$

The estimated volume ($\tilde{v}_{ik}$) of the bole of this tree from L to U based on the measurement of the kth of m cross sections equals the product of volume predicted by the taper equation and the ratio of actual to predicted cross-sectional areas at θ_k, i.e.,

$$\begin{aligned} \tilde{v}_{ik} &= \frac{A(\theta_k)}{S^*(\theta_k)} \int_L^U S^*(h)\mathrm{d}h \\ &= \frac{A(\theta_k)}{S^*(\theta_k)} T_M \qquad k = 1, 2, \ldots, m \end{aligned} \tag{14}$$

A combined estimate of the bole volume for tree i ($\tilde{v}_i$) is calculated by averaging the m separate estimates:

$$\tilde{v}_i = \frac{1}{n} \sum_{k=1}^{m} \tilde{v}_{ik} \tag{15}$$

The estimated variance of $\tilde{v}_i$ is:

$$\text{var}(\tilde{v}_i) = \frac{1}{m(m-1)} \sum_{k=1}^{m} (\tilde{v}_{ik} - \tilde{v}_i)^2 \qquad m \geq 2 \tag{16}$$

If tree i is the jth of n trees selected in the first-stage sampling, then an estimate of the total bole volume ($\tilde{V}_j$) of the survey plot is:

$$\tilde{V}_j = C_N \frac{\tilde{v}_i}{y_i} \qquad j = 1, 2, \ldots, n \tag{17}$$

The n separate estimates are averaged to give one combined estimate ($\tilde{V}$) for the plot:

$$\tilde{V} = \frac{1}{n} \sum_{j=1}^{n} \tilde{V}_j \tag{18}$$

with estimated variance:

$$\text{var}(\tilde{V}) = \frac{1}{n(n-1)} \sum_{j=1}^{n} (\tilde{V}_j - \tilde{V})^2 \qquad n \geq 2 \tag{19}$$

Note that estimated variances for the first-stage trees are not needed to calculate the estimated variance of the total volume of the plot with eqn (19). Therefore, it is not necessary to select more than one cross section per tree in the second-stage importance sampling to calculate the variance of the plot volume.

Species Breakdowns

A breakdown of volumes and variances by species can be obtained by sampling each species on the plot separately. Our computer program divides the tree list for the survey plot into tree lists for each species if the user indicates that species breakdowns are needed. The computer selects a user-specified number of sample trees at random from each species list with probability proportional to predicted size and then proceeds to direct the importance sampling of each of these trees. The unbiased estimates of the volumes and variances for the individual species obtained from this sampling are added together to produce the totals for the plot. An example of a plot sampling, by species, is presented in the appendix. The standard errors in this example are the square roots of the appropriate variances.

DISCUSSION

In order to use our computer program to sample for plot volume, the user must specify the sample sizes. The appropriate sizes of the first- and second-stage samples, n and m, will depend on the objectives of the survey. To gain some insight into whether we should emphasize n or m, we can look at the components of variance. The true variance of $\tilde{V}$, our unbiased estimate of plot volume, is:

$$\text{Var}(\tilde{V}) = \frac{\sigma_b^2}{n} + \frac{\overline{\sigma}_w^2}{mn} \tag{20}$$

where

$$\sigma_b^2 = \sum_{i=1}^{N} p_i \left(\frac{v_i}{p_i} - V \right)^2 \tag{21}$$

$$\overline{\sigma}_w^2 = \sum_{i=1}^{N} p_i \int_{L_i}^{U_i} q_i(h) \left(\frac{A_i(h)}{p_i\, q_i(h)} - \frac{v_i}{p_i} \right)^2 \mathrm{d}h \tag{22}$$

$$p_i = \frac{y_i}{C_N} \tag{23}$$

$$q_i(h) = \frac{S^*(h)}{T_M} \qquad \text{for tree } i \tag{24}$$

The quantities σ_b^2 and $\overline{\sigma}_w^2$, respectively, are the between- and within-tree components of variance, v_i is the true volume of tree i, and V is the true volume of the plot. Increasing n, the first-stage sample size, reduces the magnitude of both components of variance, whereas an increase in the second-stage sample size, m, reduces only the within-tree component. Therefore, if the accuracy of the estimate is an overriding consideration, it would be best to let m equal one and make n as large as possible. On the other hand, it is cheaper to select and measure another cross section on a given tree than to select another tree, measure H, L, and U, and then select and measure a cross section. Based on an analysis of Furnival et al. (1986), if c_1 is the average cost of moving equipment (e.g., an optical dendrometer) and measuring H, L, and U on a tree, and c_2 is the average cost of measuring a cross section on the bole, then the optimal number of cross sections to include in the importance sample per tree to achieve a given variance at the least cost is:

$$m_{\text{opt}} = \sqrt{\frac{c_1 \overline{\sigma}_w^2}{c_2 \sigma_b^2}} \tag{25}$$

The utility of this formula is somewhat questionable because σ_b and $\overline{\sigma}_w$ are never known in practice, and c_2 should vary with average tree height. Until a detailed study is done which suggests otherwise, it seems advisable to select one cross section per tree and as many trees per plot as are practicable given the budget of the survey. If the objective is an estimate of an inventory for a region, it might be wise to select one or two trees per plot (or one tree per species for each species on a plot) and concentrate on sampling as many plots as possible under the assumption that the variance of the estimate for the region will be dominated by the between-plots component of variance.

An instrument to measure upper tree diameters (or a tree climber) is essential for implementing importance sampling. Because the cross-sectional areas of tree boles decrease acropetally, most of the heights of the cross sections selected by importance sampling will be in the lower half of the bole. However, upper bole measurements will be required. Previously, we suggested a Finnish or pane-tree caliper mounted on a telescopic or sectional pole as a suitable instrument for importance sampling (Gregoire *et al.* 1986). Some boles may be too high for poles, and optical dendrometry would seem to be the only option.

REFERENCES

Furnival, G. M., Valentine H. T., and Gregoire, T. G., 1986: Estimation of log volume by importance sampling. Forest Science, vol. 32, pp. 1073-1078.

Gregoire, T. G., Valentine, H. T., and Furnival, G. M., 1986: Estimation of bole volume by importance sampling. Canadian Journal of Forest Research, vol. 16, pp. 554-557.

Gregoire, T. G., Valentine, H. T., and Furnival, G. M., 1987: Sampling methods for estimating stem volume and volume increment. Forest Ecology and Management, vol. 21, pp. 311-323.

Valentine, H. T., Tritton, L. M., and Furnival, G. M., 1984: Subsampling trees for biomass, volume, or mineral content. Forest Science, vol. 30, pp. 673-681.

Valentine, H. T., Gregoire, T. G., and Furnival, G. M., 1987: Unbiased estimation of total tree weight by three-stage sampling with probability proportional to size. pp. 129-132. *In*: Estimating tree biomass regressions and their error. USDA Forest Service General Technical Report NE-GTR-117. 303 p.

White, J., 1981: The allometric interpretation of the self-thinning rule. Journal of theoretical Biology, vol. 89, pp. 475-500.

APPENDIX

The following is an example of a two-stage sampling with the computer program. Prompts and output by the computer are in lightface type. The user's responses, all of which follow question marks, are in boldface type.

```
Random number seed (-32768 to 32767)?  44

Write output to file: [0 - NO; 1 - YES]? 1
What file?   A:EXAMPLE.DAT

Codes entered: [0; 1 - ID; 2 - SPECIES; 3 - ID, SPECIES]?  3

Enter ID CODE, SPECIES CODE, DBH (cm) of tree 1
?  A1,RED OAK,26.9

[1 - next tree; 2 - select sample; 3 - redo tree; 4 - stop]?  1
Enter ID CODE, SPECIES CODE, DBH (cm) of tree 2
?  101,HEMLOCK,28.4

[1 - next tree; 2 - select sample; 3 - redo tree; 4 - stop]?  1
Enter ID CODE, SPECIES CODE, DBH (cm) of tree 3
?  102,HEMLOCK,37.8

[1 - next tree; 2 - select sample; 3 - redo tree; 4 - stop]?  1
Enter ID CODE, SPECIES CODE, DBH (cm) of tree 4
?  A2,RED OAK,22.8

[1 - next tree; 2 - select sample; 3 - redo tree; 4 - stop]?  1
Enter ID CODE, SPECIES CODE, DBH (cm) of tree 5
?  A3,RED OAK,37.8

[1 - next tree; 2 - select sample; 3 - redo tree; 4 - stop]?  1
Enter ID CODE, SPECIES CODE, DBH (cm) of tree 6
?  103,HEMLOCK,33.7

[1 - next tree; 2 - select sample; 3 - redo tree; 4 - stop]?  2

Sample species separately: [0 - NO; 1 - YES]?  1

Number of trees of species RED OAK = 3

Enter sample sizes: n, m?  2,1
[1 - proceed; 2 - redo; 3 - stop]?  1

Tree A1 is selected with prob .2498685

[1 - enter heights; 2 - enter %slopes]?  1

Enter heights (m): H, L, U?  18.3,.3,12.
[1 - proceed; 2 - redo]?  1

Enter DIA (cm) at height 1.010397 m?  27.9
[1 - proceed; 2 - redo]?  1

Est. volume of tree A1 = .5027917     SE = na
```

```
Tree A3 is selected with prob .5848714

[1 - enter heights; 2 - enter %slopes]?  1

Enter heights (m): H, L, U?  19.2,.3,15.
[1 - proceed; 2 - redo]?  1

Enter DIA (cm) at height 2.682605 m?  36.0
[1 - proceed; 2 - redo]?  1

Est. volume of tree A3 = 1.04643      SE = na

Est. total volume of RED OAK = 1.900694      SE = .1115325

Number of trees of species HEMLOCK = 3

Enter sample sizes: n, m?  2,1
[1 - proceed; 2 - redo; 3 - stop]?  1

Tree 102 is selected with prob .4464723

[1 - enter heights; 2 - %slopes]?  2

Enter: distance to tree (m)?  20.
Enter: % to top; % to bottom (- if < 0)?  83.,-10
Enter: % to U; % to L (101 if .3 m)?  64.,101

Height = 18.6, L = .3, U = 14.8
[1 - proceed; 2 - redo]?  1

Enter DIA (cm) at height 7.661625 m (% = 28.30812 at 20 m)?  29.1
[1 - proceed; 2 - redo]?  1

Est. volume of tree 102 = .9745099      SE = na

Tree 102 is selected with prob .4464723

Enter DIA (cm) at height 8.438168 m (% = 32.19084 at 20 m)?  28.2
[1 - proceed; 2 - redo]?  1

Est. volume of tree 102 = .9844489      SE = na

Est. total volume of HEMLOCK = 2.193819      SE = 1.113453E-02

Species        Volume      Standard Error

RED OAK         1.900694     .1115325
HEMLOCK         2.193819     1.113453E-02
PLOT            4.094513     .1120869
```

Monitoring Forests Through Succesive Inventories

Moderator: James T. Bones

Charles T. Scott, An Overview of Fixed Versus Variable-Radius Plots for Successive Inventories

Michael Koehl, Designing a New National Forest Inventory for Switzerland

John R. Mills, Developing ATLAS Growth Parameters From Forest Inventory Plots

Richard Birdsey and Hans Schreuder, Forest Change - An Assessment of Nationwide Forest Inventory Methods

Charles T. Scott and James Alegria, Fixed-Versus Variable-Radius Plots for Change Estimation

William H. McWilliams, Assessing Pine Regeneration for the South Central United States

James M. Hill, Joshua L. Royte, Jil M. Swearingen, Keith Van Ness, William Warncke, and John E. Hench, A Permanent Plot System Designed for Making Natural Resources Management Decisions in a Rapidly Urbanizing County

Piermaria Corona and Agostino Ferrara, INFOR_1: An Analysis Program for Permanent Forest Inventories With Systematic Sampling

Susan L. Stout, Inventory to Support Silvicultural Decisions

M.R. Carroll, Funding, Pests and Hardwood Log Grades: User-Driven Cooperative Projects in the North Central States

Section 3

An Overview of Fixed Versus Variable-Radius Plots for Successive Inventories

Charles T. Scott
Research Forester
USDA Forest Service
Northeastern Forest Experiment Station
5 Radnor Corp. Center, #200
Radnor, PA 19087

ABSTRACT

Since Bitterlich introduced point or variable-radius sampling in 1947, many investigators have compared it with fixed-area sampling for estimation of current attributes. A partial review of the literature that compares the two methods is given for successive or continuous forest inventories. The sampling methods are described in the areas of field implementation, components of change estimation, comparison methods used, and efficiency for both current and change estimates. Point sampling is good for current attributes which are related to tree size, but this advantage diminishes for change estimation.

INTRODUCTION

Forest inventories are usually based on one of two sampling systems, fixed-area or variable-radius sampling. Fixed-area samples are often circular (fixed-radius) and have been used for decades, particularly with Continuous Forest Inventories (Ware and Cunia 1962). Variable-radius or point sampling was introduced by Bitterlich (1947). Point sampling is very efficient for estimating basal area per acre. Measuring tree diameter was not required---only a tree count using a prism or angle gauge (Grosenbaugh 1952). Investigators have used many different methods to compare point sampling against the time-honored fixed-area sampling rule to ensure that the estimates were both accurate and cost-effective.

Initially, point sampling was used only for quick, temporary surveys without monumenting each plot. Over time the need for estimates of change between successive surveys became important. In some cases, the only plots available for remeasurement were point samples. Many articles have appeared in the literature on how to estimate growth from point samples. Eventually, a few investigators compared fixed- and variable-radius sampling for change estimation. This article presents an overview of this history, but focuses on comparing fixed plots and point samples for successive surveys. No attempt has been made to give a complete bibliography of the subject.

PRACTICAL ASPECTS

Much of the initial appeal of point sampling was that sample trees need not be visited to compute precise estimates of basal area and volume (Grosenbaugh and Stover 1957). For successive surveys, components of change (ingrowth, accretion, mortality, and removals) can only be estimated by monumenting the plot and identifying the location of each tree (Beers 1962). Thus, a primary benefit of point sampling is lost for successive surveys.

How to handle borderline trees with point sampling has long been a concern. Zeide and Troxell (1979) and Wiant et al. (1984) investigated concerns over missing trees as the distance between the sample point and tree increases. Grosenbaugh (1952) and Solomon (1975) also found underestimates using a small basal area factor (BAF). Brush and trees obscure the observer's vision. Point samples are more affected than fixed plots, so Wiant et al. and Zeide and Troxell recommended using large BAF's to minimize the number of missed trees.

Another disturbing factor is that the effective plot area for a tree changes as it grows, thus changing the number of trees per acre the tree represents (Beers and Miller 1964).

COMPONENTS OF CHANGE ESTIMATION

The components of change usually include ingrowth, accretion (survivor growth), mortality, removals (cut), and net change. These components can be applied to a variety of attributes, such as number of stems, basal area, or volume per acre. The methods for estimating these are simple for fixed plots. When Beers (1962) introduced methods for estimating components of change from point samples, he noted some "complications." The primary problems were: 1) not all new trees are ingrowth, some can be "nongrowth" trees, 2) the number of trees (expansion factor) changes over time for survivor trees, and 3) the components of change did not necessarily sum to the difference between the two survey estimates (see Roesch et al. 1989).

"Nongrowth" trees were above the merchantability limit at the previous survey and were not sampled with the prism, but are large enough now to be sampled with the prism. Various authors have tried different methods of handling these new trees and have called them either "ongrowth" or "nongrowth" trees (Beers 1962, Beers and Miller 1964, Nishizawa 1967, and Martin 1982). Van Deusen et al. (1986) used diameter growth models to distinguish between nongrowth and true ingrowth trees and included their current volume as survivor growth. Roesch et al. (1989) included the nongrowth component in a more natural way by including only their estimated growth.

The second problem is one of choosing an expansion factor to apply to survivor trees. One suggestion was to use only the previous expansion factor (Myers and Beers 1968). Van Deusen et al. (1986) suggested using the current expansion factor for the current attribute and the previous one for the previous attribute, then taking the difference. This can lead to negative growth estimates for survivor trees, but is offset by including all of the current volume of nongrowth trees---not just the growth. Other approaches have been tried by Martin (1982) and Iles and Beers (1983). Roesch et al. (1989) used only the current expansion factor for both previous and current attributes for both survivor and nongrowth trees. However, the additivity criterion was not satisfied without improving the estimate at Time 1 based on information recorded at Time 2.

METHODS OF COMPARING SAMPLE DESIGNS

The two primary methods of comparing the sampling designs are to compute the **relative precision** (ignoring costs) or the **relative efficiency**. Many studies have reported the relative precision of particular plot designs that were tried in the field (Grosenbaugh and Stover 1957, Solomon 1975). The primary benefit of these studies was to verify that point samples gave accurate results. They provided little indication of which sampling rule was most cost-effective.

Relative efficiency can be computed in one of two ways: taking the ratio of variances for a fixed cost, or taking the ratio of costs for fixed precision. When computing relative efficiency, a cost function is implied. Many investigators have simply observed costs for specific plot designs and made direct comparisons (Avery and Newton 1965, Barrett and Carter 1968). Stage (1958) presented a graph relating fixed plot size and BAF for individual tree diameters. Banyard (1976) fixed the cost by choosing fixed plot and point sample sizes so that the number of trees were equal. Several investigators used detailed cost functions to closely approximate field conditions (Nyyssonen and Kilkki 1965, Nichols 1980, Scott 1981a, Scott et al. 1983). When these cost functions were associated with a model for plot variability (precision), optimal plot designs could be found. Either the cost was minimized for a specified precision level or variance was minimized for fixed cost.

Cost Functions

Investigators have used a variety of cost functions. Functions for single plots usually involve overhead costs and the cost per plot times the number of plots. Other studies included the cost of travel between sample plots as a function of the density of plots in the population (Jessen 1942, Humphreys 1979). In the case of cluster sampling, Hansen et al. (1953) present a simple cost function. In forestry, Anagnostopoulos (1966), Nyyssonen et al. (1971), O'Regan et al. (1973), Nichols (1980), and Scott (1981a) presented detailed cost functions for cluster sampling. The approach of Scott (1981a) was to model actual survey costs and the variation between clusters as functions of the cluster design variables. The cluster design variables were the number of plots per cluster (m), the average distance between plots (d), and the size of each plot (z).

Precision Models

Perhaps the first model to relate variability to plot design characteristics was introduced by Smith (1938). This model form has been used frequently in forestry applications (Freese 1961, Ziede 1980, Wiant and Yandle 1980). Nyyssonen et al. (1971) and Scott (1981a) expanded Smith's model form to other cluster design variables.

Optimization

Many investigators have used optimization techniques to identify cost-effective sampling designs. Perhaps the most direct method is to use theoretical methods. O'Regan and Arvanitis (1966) computed the relative efficiency of fixed plots and point samples after analytically minimizing cost subject to constraints on precision models. Sukwong et al. (1971), Matern (1972), Oderwald (1975, 1979, 1981), Ranneby (1980), and Martin (1983) used similar mathematical techniques but assumed cost was constant. However, they introduced characteristics of the population, such as spatial distribution, into the precision model in an attempt to generalize the results. Gambill et al. (1985) did not compare the two sampling methods, but presented tables of the optimal fixed plot and point sample sizes for varying tract sizes, coefficients of variation, and measurement times.

When the cost or precision models become too complicated to derive mathematical optima, numerical optimization techniques such as linear programming have been used (Hazard and Promnitz 1974, Hazard 1976, Marshall and Nautiyal 1980, Morris 1981). Others have used Monte Carlo simulation techniques. A real or simulated stem map of a forest can be used to test many sampling designs without performing additional field work (Kulow 1966; Arvanitis and O'Regan 1967, 1972; Kaltenberg 1978).

COMPARISONS FOR CURRENT ESTIMATES

Soon after Grosenbaugh (1952) introduced point sampling to the United States forestry literature, many investigators began comparing it with fixed plot sampling for estimating current values. Using theoretical methods, Sukwong et al. (1971) and Matern (1972) state that fixed plot sampling must always sample more trees than point samples to yield the same precision. They provide general solutions for comparing fixed plots and point samples when costs are ignored by including the effects of clustering and diameter distributions. Oderwald (1981) states that point samples perform best when there is a wide range of diameters.

John (1964) found that point samples are better for estimation of cover type. Lindsey (1956) found point samples insufficient for density estimation, so recommended augmenting them with quadrats. Lindsey et al. (1958) found point samples to be more efficient than fixed plots when estimating both density and basal area. Nyyssonen and Kilkki (1965) compared several fixed plot and point sample sizes with the 10-point prism cluster often used by the USDA Forest Service (Scott 1981b). Fixed-plot sizes of 0.1--0.125 ha were the most efficient of the designs considered for volume estimation over a range of conditions. O'Regan and Arvanitis (1966) state that fixed plots were most efficient for estimating number of trees and point samples are most efficient for estimating basal area. No conclusion was given in the case where both attributes are of interest.

Joshi and Tomar (1976) found point samples to be more cost-effective even though point sampling required more 1 m^2/ha BAF plots than 0.05--0.125 ha fixed plots. Banyard (1976) found point samples to be 2.78 times more efficient than fixed plots for current basal area estimates in tropical rain forests. He also noted that the stand table for all but the smallest diameter class was better estimated with point samples due to the higher rate of sampling of large trees. Kulow (1966) recommended a 10 BAF plot over fixed plots for basal-area estimation. Barrett and Carter (1968) found 20-BAF plots to be 67--98 percent more efficient than 1/10-acre plots for volume estimation. Avery and Newton (1965) found 1/10-acre and 10-BAF plots to perform equally well for bottomland hardwoods in Georgia, but 1/10- or 1/20-acre plots were more efficient for natural pine stands.

For northern Ontario, Payandeh (1974) found that 1) most coniferous forests are highly clustered, 2) most hardwood forests are randomly distributed, and 3) uniformly distributed natural stands are rare. Martin (1983) states that if the attribute of interest is independent of, or negatively correlated with, basal area, then fixed plots are always more precise (see also Oderwald 1975, 1979, 1981). He provides formulas for a critical coefficient of correlation under random or clustered spatial distributions. Correlations with basal area greater than the critical value indicate that point samples are more precise. Oderwald (1981) states that "point sampling is more precise than plot sampling for the same effort in the clumped and random spatial patterns, but plot sampling is more precise in the square lattice pattern." Applying Payandeh's results, we would expect that point sampling is more precise for natural stands, and plot sampling is more precise for planted stands. However, the study by Avery and Newton (1965) does not support this.

COMPARISONS FOR CHANGE ESTIMATES

The work of Martin (1983) focused on the relative precision of growth estimators, but his results apply equally well to the estimation of current values discussed in the preceeding section. Thus the relative precision of fixed plots to point samples depends on the spatial distribution and the correlation of the attribute of interest to basal area.

Arvanitis and O'Regan (1967) used simulation techniques to compare fixed plots and point samples for the estimation of current number of trees, number of ingrowth trees, current basal area, and gross basal-area growth (ingrowth plus accretion). If we assume that measurement costs were the same for trees sampled on fixed plots versus point samples, fixed plots were more efficient for all attributes except current basal area. Arvanitis and O'Regan (1967, p. 149) found gross basal-area growth to be "more highly correlated with number of trees than with basal area itself when estimates are based on point sampling." Elsewhere in these proceedings, Scott and Alegria found that basal area of ingrowth trees was more efficiently estimated on fixed plots, but basal-area accretion was more efficiently estimated on point samples. Thus the results of Arvanitis and O'Regan may be due in part to using gross basal-area growth rather the basal-area accretion.

Ranneby (1980) extended Matern's (1972) work to the estimation of basal-area growth. He states that the advantage of point samples is less when estimating basal-area growth versus current basal area.

The decrease is substantial if no cutting has occurred. He states that the efficiency of point samples for estimating current basal area can also hold for estimating removals. The empirical results of Scott and Alegria agree in part with Ranneby's statements. Scott and Alegria found removals to be efficiently estimated using point samples, but the relative efficiency of basal-area growth was as least as large as for current basal area.

CONCLUSIONS AND SUMMARY

The practical problems with using point sampling for permanent plots in successive inventories largely have been overcome through the use of appropriate and efficient growth estimators. Many comparisons of fixed- and variable-radius sampling have been performed while ignoring costs. Many others have addressed only current basal area estimation. More work needs to done using mathematical and numerical optimization techniques to find cost-effective designs for a variety of current and change attributes across a range of diameter and spatial distributions.

In general, fixed plots were found to be more efficient for estimation of the current number of stems. Point samples were more efficient for current basal area and volume. The efficiency advantage of point samples is reduced when estimating accretion in basal area and volume. Fixed plots are likely to be more efficient for estimation of the change in number of stems and of ingrowth basal area and volume. As with estimation of current values, the choice between fixed plots and point samples for estimating change depends on the relative importance of estimating number of trees versus basal area or volume.

LITERATURE CITED

Anagnostopoulos, C.J., 1966: The most efficient number of elements and distance between them in a cluster sampling unit. M.S. thesis, State University of New York, Syracuse, 72 pp.

Arvanitis, L.G., O'Regan, W.G., 1967: Computer simulation and economic efficiency in forest sampling. Hilgardia, vol. 38, pp. 133--164.

Arvanitis, L.G., O'Regan, W.G., 1972: Cluster or satellite sampling in forestry: a Monte Carlo computer simulation study. Proceedings Third Conference IUFRO Advisory Group of Forest Statisticians, Jouy-en-Josas, France. pp. 191--205.

Avery, Gene and Newton, R., 1965. Plot sizes for timber cruising in Georgia. Journal of Forestry, vol. 63, no. 12, pp. 930--932.

Banyard, S.G., 1976: A comparison between point sampling and plot sampling in tropical rain forest based on a concept of the equivalent relascope plot size. Commonwealth Forestry Review, vol. 54, pp. 312--320.

Barrett, J.P. and Carter, J.F., 1968: Comparing methods of sampling a white pine stand. Journal of Forestry, vol. 66, no. 11, pp. 861--863.

Beers, T.W., 1962: Components of forest growth. Journal of Forestry, vol. 60, no. 4, pp. 245--248.

Beers, T.W. and Miller, C.I., 1964. Point sampling: research results, theory, and applications. Purdue University Research Bulletin 786. 56 pp.

Bitterlich, W., 1947: Die winkezahlmessung. Allg. Forst-u. Hozwirtsch. Ztg., vol. 58, pp. 94--96.

Freese, F., 1961: Relation of plot size to variability: an approximation. Journal of Forestry, vol. 59, p. 679.

Gambill, C.W., Wiant, H.V., Jr. and Yandle, D.O., 1985: Optimum plot size and BAF. Forest Science vol, 31, no. 3, pp. 587--594.

Grosenbaugh, L.R., 1952: Plotless timber estimates---new, fast, easy. Journal of Forestry, vol. 50, no. 1, pp. 32--37.

Grosenbaugh, L.R. and Stover, W.F., 1957: Point sampling compared with plot sampling in southeast Texas. Forest Science, vol. 3, no. 1, pp. 2--14.

Hansen, M.H., Hurwitz, W.N., and Madow, W.G., 1953: Sample Survey Methods and Theory, Vol. I., John Wiley and Sons, NY, 638 pp.

Hazard, J.W., 1976: Optimization in multi-resource inventories. Proceedings Inventory Design and Analysis. W.E. Frayer, G.B. Hartman, and D.R. Bower, editors. Society of American Foresters workshop, pp. 170--180.

Hazard, J.W. and Promnitz, L.C., 1974. Design of successive forest inventories: optimization by convex mathematical programming. Forest Science, vol. 20, pp. 117--127.

Humphreys, C.P., 1979: The cost of sample survey designs. Proceedings American Statistical Association, Survey Research Methods Section, pp. 395--400.

Iles, K. and Beers, T.W., 1983: Growth information from variable plot sampling. Proceedings Renewable Resource Inventories for Monitoring Changes and Trends, College of Forestry, Oregon State University, Corvallis, Society of American Foresters 83-14. pp. 693--694.

Jessen, R.J., 1942: Statistical investigation of a sample survey for obtaining farm facts. Iowa Agricultural Experiment Station Research Bulletin 304, 104 pp.

John, H.H., 1964: Comparative precision levels in forest inventory estimating using sampling with probability proportional to frequency and sampling with probability proportional to size. Ph.D. dissertation, University of Minnesota, St. Paul. 76 pp.

Joshi, S.C. and Tomar, M.S., 1976: Investigations on plot size vis-a-vis point sampling: a pilot study. Indian Forester, vol. 102, no. 10, pp. 701--711.

Kaltenberg, M.C., 1978: Evaluation of regeneration sampling methods: a Monte Carlo analysis using simulated stands. State of Washington, Department of Natural Resources, Report 39, 50 pp.

Kulow, D.L., 1966: Comparison of forest sampling designs. Journal of Forestry, vol. 64, no. 7, pp. 469--474.

Lindsey, A.A., 1956: Sampling methods and community attributes in forest ecology. Forest Science, vol. 2, no. 1, pp. 287--296.

Lindsey, A.A., Barton, J.D., Jr., and Miles, S.R., 1958: Field efficiencies of forest sampling methods. Ecology, vol. 39, no. 3, pp. 428--444.

Marshall, P.L. and Nautiyal, J.C., 1980: Optimal sampling size in multivariate forest inventories: a programming procedure. Canadian Journal of Forest Research, vol. 10, no. 4, pp. 579--585.

Martin, G.L., 1982: A method for estimating ingrowth on permanent horizontal sample points. Forest Science, vol. 28, no. 1, pp. 110--114.

Martin, G.L., 1983: The relative efficiency of some forest growth estimators. Biometrics, vol. 39, no. 3, pp. 639--650.

Matern, B., 1972: The precision of basal area estimates. Forest Science, vol. 18, no. 2, pp. 123--125.

Morris, M.J., 1981: Plot configurations for multiresource purposes in the West. Proceedings In-Place Resource Inventories: Principles and Practices. Orono, ME, Society of American Foresters 82-2, pp. 375--378.

Myers, C.C. and Beers, T.W., 1968: Point sampling for forest growth estimation. Journal of Forestry, vol. 66, no. 12, pp. 927--929.

Nichols, J.D., 1980: Intra-cluster correlation in sample design. Proceedings Arid Land Resource Inventories: Developing cost-efficient methods. USDA Forest Service General Technical Report WO-28, pp. 142--147.

Nishizawa, M., 1967: Growth estimation by the angle method. Proceedings 14th Congress Internation Union of Forest Research Organizations, Munich, Part VI, Section 25, pp. 410--425.

Nyyssonen, A. and Kilkki, P., 1965: Sampling a stand in forest survey. Acta Forestalia Fennica, vol. 79, no. 4, 20 pp.

Nyyssonen, A., Roiko-Jokela, P., and Kilkki, P., 1971: Studies on improvement of the efficiency of systematic sampling in forest inventory. Acta Forestalia Fennica, vol. 116, 26 pp.

Oderwald, R.G., 1975: The effect of spatial distribution on the efficiency of a class of probability proportional to size sampling estimators. Unpubl. Ph.D. dissertation, University of Georgia, Athens. 63 pp.

Oderwald, R.G., 1979: Comparison of the efficiencies of point and plot sample basal area estimates in differing stand situations. Proceedings Forest Resource Inventories, Ft. Collins, CO, vol. I, pp. 250--260.

Oderwald, R.G., 1981: Comparison of point and plot sampling basal area estimators. Forest Science, vol. 27, no. 1, pp. 42--48.

O'Regan, W.G. and Arvanitis, L.G., 1966: Cost-effectiveness in forest sampling. Forest Science, vol. 12, no. 4, pp. 406--414.

O'Regan, W.G., Seegrist, D.W. and Hubbard, R.L., 1973: Computer simulation and vegetation sampling. Journal of Wildlife Management, vol. 37, pp. 217--222.

Payandeh, B., 1974: Spatial pattern of trees in the major forest types of northern Ontario. Canadian Journal of Forest Research, vol. 4, pp. 8--14.

Ranneby, B., 1980: The precision of the estimates of the growth in basal area. Swedish University of Agricultural Sciences, Section of Forest Biometry, Report no. 19, 25 pp.

Roesch, F.A., Jr., Green, E.J. and Scott, C.T., 1989: New compatible estimators for survivor growth and ingrowth from remeasured horizontal point samples. Forest Science, vol. 35, no. 2, pp. 281--293.

Scott, C.T., 1981a: Design of optimal two-stage multiresource surveys. Ph.D. dissertation, University of Minnesota, St. Paul, 138 pp.

Scott, C.T., 1981b: Plot configurations in the East for multiresource purposes. Proceedings In-Place Resource Inventories: Principles and Practices. Orono, ME, Society of American Foresters 82-2, pp. 379--382.

Scott, C.T., Ek, A.R., and Zeisler, T.R., 1983: Optimal spacing of plots comprising cluster in extensive forest inventories. Proceedings Renewable Resource Inventories for Monitoring Changes and Trends, College of Forestry, Oregon State University, Corvallis, Society of American Foresters 83-14, pp. 707--710.

Smith, H.F., 1938: An empirical law describing heterogeneity in the yields of agricultural crops. Journal of Agricultural Science, vol. 28, no. 1, pp. 1--23.

Solomon, D.S., 1975: A test of point-sampling in northern hardwoods. USDA, Forest Service, Research Note, NE-215, 2 pp.

Stage, A.R., 1958: An aid for comparing variable plot radius with fixed plot radius cruise designs. Journal of Forestry, vol. 56, p. 593.

Sukwong, S., Frayer, W.E. and Mogren, E.W., 1971: Generalized comparisons of the precision of fixed-radius and variable-radius plots for basal-area estimates. Forest Science, vol. 17, no. 2, pp. 263--271.

Van Deusen, P.C., Dell, T.R. and Thomas, C.E., 1986: Volume growth estimation from permanent horizontal points. Forest Science, vol. 32, pp. 415--422.

Ware K.D. and Cunia, T., 1962: Continuous forest inventory with partial replacement of samples. Forest Science Monograph 3, 40 pp.

Wiant, H.V., Jr. and Yandle, D.O., 1980: Optimum plot size for cruising sawtimber in eastern forests. Journal of Forestry, vol. 78, pp. 642--643.

Wiant, H.V., Jr., Yandle, D.O., and Andreas, R., 1984: Is BAF 10 a good choice for point sampling? Northern Journal of Applied Forestry, vol. 1, pp. 23--24.

Zeide, B., 1980: Plot size optimization. Forest Science, vol. 26, no. 2, pp. 251--257.

Zeide, B. and Troxell, J.K., 1979: Selection of the proper metric basal area factor for Appalachian mixed hardwoods. Proceedings Forest Resource Inventories, Ft. Collins, CO, vol. I, pp. 261--269.

Designing a New National Forest Inventory for Switzerland

Michael Koehl
Swiss Federal Institute for Forest, Snow and Landscape
Dept.NFI
CH-8903 Birmensdorf, Switzerland

ABSTRACT

The first Swiss National Forest Inventory was conducted as a systematic survey over a 1 x 1 km grid, and since the aim from the start was to establish a permanent inventory all the 10 975 sample plots were permanently marked. The first inventory was carried out between 1983 and 1985, and the findings were published in 1988. The second inventory is planned to begin in 1992. The main parameters to be determined are standing volume, increment, drain, and site conditions. Among other considerations, the application of aerial photography, stratification, and SPR as a concept for a permanent inventory are being studied in the further development of the design.

1st National Inventory

The first national inventory was conducted between 1983 and 1985. The main interpretation of the data was completed in 1988, when the results were published (Bachofen et al., 1988). Special evaluations for various interested parties such as cantonal Forest Services, free-lance forest engineers, and scientific institutes are still under way. The aim of the first NFI was to determine the condition of forests in Switzerland on a large-scale basis so as to furnish a basis for decision-making in forest policy.

The main emphasis was on the collection of data concerning standing volume, forest area, assortments and quality of timber, site characteristics, and accessibility.

The data were collected from four different sources:
- aerial photographs (black and white, 1:23000)
- topographic maps
- field samples
- questionnaires circulated among the regional Forest Services.

A 1 x 1 km grid was laid out over the entire country, and apart from those areas occupied by water or above 2500 m, i.e., above the timberline, all intersection points on aerial photographs were evaluated with an analytical plotter. Aerial photography was employed for two purposes:
- first, to determine the area of forest (with distinction between the three groups forest, shrub growth, and non-forest);
- secondly, to select suitable reference points for the location of the centres of sample plots.

All points classified on aerial photographs as forest were also terrestrially surveyed and permanently marked. This means that the sampling areas were also systematically distributed over a 1 x 1 km network covering the entire country. It is possible for each canton, through reducing the scale of the network, to obtain data on smaller units. Given the coordinates determined from the aerial photographs, the centres of the sample plots can be determined with the help of reference points. The actual location can be established with the help of simple instruments (compass and metre rule) and a pocket calculator programme.

Each sample unit comprises not only two concentric circular plots but also an interpretation area of 50 x 50 m. On the latter, data on area and stand conditions were collected. On the smaller circular plots, with an area of 200 m2, all trees with a DBH over 12 cm, and on the larger plots, with an area of 500 m2, all those with a DBH above 36 cm were registered. All trees with an azimuth of 0 to 150 or with a DBH of 60 cm or more were taken as sample tariff trees. For these trees, both the height and the diameter at 7 m (measured with Finnish calipers) were also determined (Zingg, 1988).

Some 8% of the total area sampled was subjected to a second survey to check the data. Apart from the countable and measurable data, some 60 different qualitative parameters were recorded, the results being compared with those of the original survey. The mean errors were found to be 0.4 cm for DBH, 1.2 cm for D7, and 1.1 m for height.

The personnel required for the first NFI is shown in Fig. 1. It should be borne in mind, however, that this does not include the planning period from 1973 to 1982. Costs for the first NFI totalled some SFr.7.5 million (Sutter et al., 1987).

2nd National Forest Inventory

Taking the findings of the first NFI, the state of health of the forest in Switzerland, and weather disasters in the Alps as a basis, the second NFI should aim at meeting present needs for information. The definition of forest functions, especially in terms of protection, must be one of the major aims, and equally, the study of site conditions must be intensified.

Nevertheless much emphasis must still be laid on determining standing volume, increment, drain, and area. These parameters are being employed in optimizing the sample design for the second inventory.

At present, studies for the design of the second inventory are concentrating on three points:

- the application of aerial photography in multi-phase inventories;
- stratified sampling taking the five productive regions as strata;
- sampling with partial replacement as the most suitable method of estimating increment.

Aerial Photography.In the second NFI,as a two-phase inventory, the application of aerial photography is to be extended. Poor accessibility and difficult terrain in the Alps and on their southern slopes render terrestrial surveys costly, so that aerial photography offers an attractive alternative in these regions. Black-and-white aerial photographs at a scale of 1:23 000 can be obtained from the Federal Office of Topography, which conducts aerial surveys of the entire country in a six-year rota. The dates and scale of the photograph are naturally fixed and must be taken as pre-set conditions in a combined inventory. The southern slopes of the Alps were selected as the test area for the combined inventory, as here there is the greatest variation in standing volume, because the stand structures and forest types are heterogeneous and the altitude ranges from 196 m to 2217 m. It was here that the terrestrial surveys were most time-consuming. For the pilot inventory some 250 sample plots, for which data from the terrestrial survey of the first NFI are available, were interpreted by means of aerial photography. Subsequently, two sampling procedures were tested:

- double sampling with regression estimators
- double sampling for stratification.

The following discussion deals solely with the determination of standing volume.

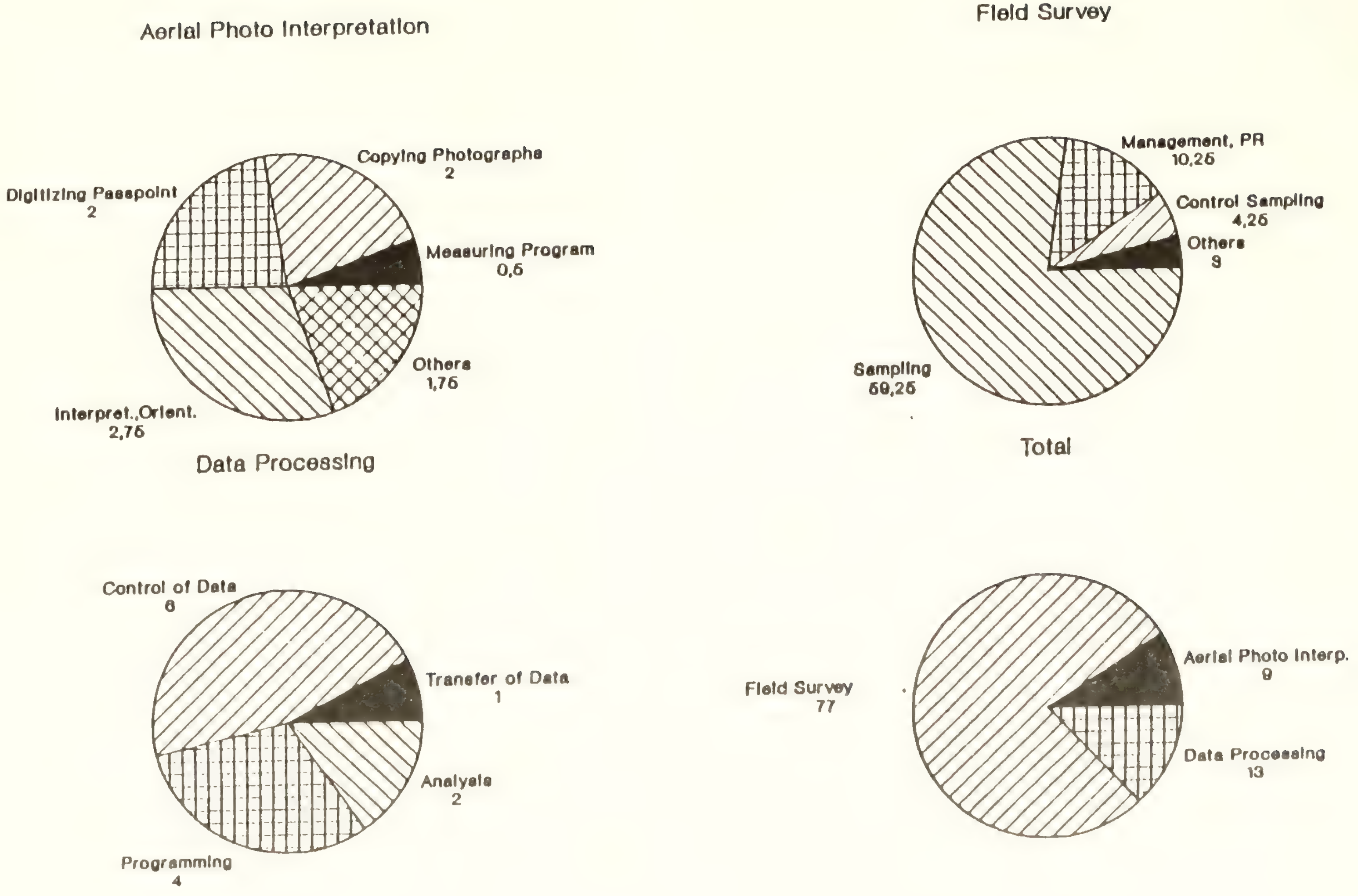


Fig.1 Time Consumption (Man Years)

Double Sampling with Regression Estimators.The first step was to compute a volume function from parameters measured on the aerial photographs. To this end, the data set was halved and the regression for one half calculated. The volume as determined in the terrestrial surveys on the basis of tariff functions was taken as the dependent variable. The independent variables (tree height as the mean of three measurements, crown length as the mean of nine measurements, density of cover, altitude from which the photograph was taken, cumulative production), all of which can be determined from the photograph, were employed to compute two volume functions:

V = f (cumulative production, tree height)
V = f (cumulative production, crown length).

These variables also proved significant in separate regression calculations for various forest types, stages of development, mixture proportions, and stand structures, though the multiple R^2 was usually below 0.4.

The other half of the data set was employed to compare the volumes determined from the terrestrial surveys and the aerial photographs. Even for sub-units such as forest type, stage of development, mixture proportion, stand structure, and degree of cover no significant differences were observed.

Once these regression calculations had been completed, 30 samples with a sample size m, i.e., between 50 and 200, were simulated at the second level in order to determine the correlation between the variable of interest (terrestrial volume) and the auxiliary variable (volume as computed from aerial photographs). Where the sample size m was smaller than 100, R^2 between the individual simulations varied greatly (0.6 to 0.05). With 100 samples or more it stabilized between 0.2 and 0.3.

As a last step, the costs for the combined inventory were compared with those of a purely terrestrial survey. In the first NFI, it took 516 min. to survey each sample plot on the southern slopes of the Alps. As the costs for the photo-interpretation of one sample plot depend on the number of sample plots per stereo model, the optimum selection had to be done by iteration.

Assuming the standard error for one-phase sampling, the optimization results in a reduction of 7 to 12% in labour, depending on the difficulties of orienting the aerial photograph.

Through optimization given the total expenditure, the combined inventory allowed a reduction in error of 3% to 4%. In calculating costs, the investments in equipment for interpretation and in programming were not taken into account; for the aerial photographs the cost of the copies but not of the actual flight coverage had to be considered. If the latter were taken into account, a comparison of costs would show the combined inventory to be disadvantageous.

The results of the pilot inventory do not agree with those of other combined inventories using double sampling with regression estimators. The main reasons for this are doubtless the difficulties due to the great variation in the forest within small areas and the scale of the aerial photographs (1:23000). At present, in order to determine the influence of the accuracy of locating the centres of the two sample plots, the centres of the terrestrial sample plots are being marked so as to be visible on the aerial photographs, so that the distance between the centres of the terrestrial and aerial sample plots can be established.

Double Sampling for Stratification. As the results of the regression sampling indicated that a combined inventory is not particularly suitable for determining standing volume in the test area, double sampling for stratification (Johnson, 1982) was investigated as an alternative. Here, instead of taking a parameter from the aerial photograph as an auxiliary variable and relating it to the variable of interest through calculating the regression, the allocation to a particular stratum was taken as the auxiliary variable.

This procedure has been in application in the national inventory of Finland since the 70's (Mattila, 1985). In the test area the most suitable basis for stratification was found to be the stage of development. This was classified in five strata: juvenile forest, polewood, young to middle-aged timber, mature timber, and mixed.

In order to determine the variance of the mean for different sample sizes in the second phase, samples were simulated, 30 simulations being conducted for each of the sample fractions 0.3, 0.5, 0.7, and 0.9 per stratum.

Table 1 shows the costs for the various sample fractions. Even with a sample fractions of only 0.7, the costs as compared with a one-phase inventory can be reduced by about 25%. As this reduction in costs involves only a slight increase in standard error (see Table 2), two-phase sampling for stratification must be regarded as more suitable than two-phase sampling with regression estimators, and will consequently be applied, at least in some parts of Switzerland, in the second NFI.

Sampling with Partial Replacement. The main aim of the NFI - large-scale monitoring of changes in the forests of Switzerland - can only be achieved through successive inventories. Consequently the NFI was designed with a view to the establishment of a permanent inventory. All the sample plots surveyed in the first NFI - somewhat more than 10 000 in number - were permanently marked so that they could be relocated. The centres were marked with aluminium stakes and the sample plot with coloured signs. For each tree, the azimuth and distance from the plot centre were recorded. Tree mapping and survey records will allow relocation even after several years.

Since (a) permanent and temporary sample plots yield equally accurate figures for standing volume; (b) only permanent sample plots provide direct information on changes in increment; (c) the degree of accuracy required in determining increment differs from that for standing volume; and (d) temporary survey plots are less expensive to survey than permanent ones, sufficient permanent plots should be available to determine increment, and, if necessary, additional temporary plots should be surveyed to guarantee the higher degree of accuracy required for the estimation of standing volume.

The concept of a permanent inventory demands that the permanent sample plots remain representative for the entire area. To ensure this, and to preclude any influence of management, the exact location of the sample plot centre should not be known to the local foresters. As in the first NFI they understood the significance of the coloured markings and the location of the sample plot centres coincided exactly with the intersections of the grid on the topographical maps, it must be assumed that the location of the sample plots is known. Further, some 700 sample plots are surveyed each year in forest damage inventories, while a third of the cantons use the NFI sample plots for their own forest damage inventories. In order to verify the representativeness of the results and establish whether any adaptation is necessary, it is absolutely essential that temporary sample plots be employed in the second NFI, some of which should be permanently marked for future use in subsequent national inventories or other purposes such as cantonal forest damage inventories.

Both in terms of accuracy and flexibility SPR appears the most suitable method for the second NFI.

Stratification. The aim of stratification is to divide a population into classes in such a way that the variance within each class is less than that of the population as a whole. Other factors to be considered include working conditions in a given area and different evaluation of the primary variables; for example, in many areas of the Alps the protective function of the forest is considered more important than timber production.

Since a stratification based on purely statistical considerations will not be accepted, the existing divisions within the country must be employed as strata. Beside the cantons, which because of their

Tab.1:Double sampling for stratification: estimated costs

Costs* per field plot : 571 min.
Costs* per photo plot[1]: 38.3 min

Sample design	number of field plots	costs(h)	number of photo plots	total costs(h)
no photo samples	236	2246	-	2246
double sampling				
s.f.= 0.3	71	676	236	827
s.f.= 0.5	118	1123	236	1274
s.f.= 0.7	165	1571	236	1722
s.f.= 0.9	212	2018	236	2169

* only variable costs

[1] relative orientation of stereo model

Tab.2 Double sampling for stratification: Costs and standard error.

Sampling design	costs (h)	S.E.	cost-difference in relation to 1-phase sampling
1-phase sampling	2246	8.94	
double sampling for stratification, sampling fraction:			
0.3	827	14.58	-63.2%
0.5	1274	11.62	-43.3%
0.7	1722	9.98	-23.3%
0.9	2169	9.07	-3.4%

differing size and purely political delimitation are unsuitable as strata, the country has been divided into five production regions, that have been used in forest statistics for some time. These units allow a rough determination of the various conditions for production, which are also reflected in the objectives of the second NFI. The standing volume in the different regions varies between 176 and 417 m^3, while the yield ranges from 0.3% on the southern slopes of the Alps and 2.1% in the Mittelland.

The cumulative production can be employed as a measure for the estimation of increment. It is defined as the maximum attainable mean increment per hectare and year under optimum stocking and management. It is estimated by means of models based on the site characteristics elevation, exposure, growth region, slope inclination, relief, and acidity of the bedrock.

As regards permanent inventories, there are certain objections to stratified samples:

- where other proportional data sets are employed, the computation of unbiased estimators may become very difficult after several surveys because the formulae may become very complicated;
- large cantons extending over more than one production region require subdivisions in evaluation;
- the production regions are not ideal as strata for all variables;
- it is uncertain whether the strata will remain constant in the course of time.

With a view to a flexible inventory design, no pre-stratification is applied. The sample plots are to be systematically surveyed without consideration of the strata limits. For the computation of estimated values, post-stratification is to be used.

Summary

The design of the second Swiss National Forest Inventory will be characterized by the transition from a pure determination of current conditions to a permanent inventory permitting the monitoring of changes.

The procedure to be used will be sampling with partial replacement. Aerial photography will be applied not primarily for large-scale surveys but will be concentrated on areas for which information on standing volume is inadequate or which are difficult of access. The application of double sampling for stratification will allow the subsequent use of satellite data.

A great deal of emphasis is being laid on flexibility of the design. Adaptation for new questions arising in the inventories of the twenty-first century must be possible and is decisive for the final selection of the inventory method.

LITERATURE

Bachofen H. et al., 1988: Schweizerisches Landesforstinventar, Ergebnisse der Erstaufnahme, EAFV,Bericht No. 305.

Johnson D.C., 1982: Theory and Application of Selected Multilevel Sampling Designs, PhD-Thesis, Colorado State University, Fort Collins, Colorado.

Mattila E., 1985: The Combined Use of Systematic Field and Photo Samples in a Large-Scale Forest Inventory in North Finland, Communicationes Instituti Forestalis Fenniae No. 131.

Sutter R. et al., 1987: The Swiss National Forest Inventory, in: FAO/ECE/FINNIDA, Forest Resource Assessment Meeting, Kotka, Finland, 26-30 October 1987.

Zingg A., 1988: Schweizerisches Landesforstinventar, Anleitung fuer die Erstaufnahme 1982-1986, EAFV, Bericht No. 304.

Developing ATLAS Growth Parameters From Forest Inventory Plots

John R. Mills
Research Forester
USDA Forest Service, Pacific Northwest Research Station
Portland, Oregon 97208

ABSTRACT

The implementation of an age-based stand projection model for the 1989 RPA timber assessment was dependent on the utilization of inventory plot data. The projection mechanism in the aggregate timberland assessment system (ATLAS) model requires yield tables and stocking adjustment equations to project stand growth. A national inventory database of plot summary records, assembled by the Forest Inventory and Analysis work units of the USDA Forest Service, was the source of data for most model inputs. Techniques were developed to generate yield tables and stocking adjustment parameters from the growth variables in the database. The techniques assume the latest forest measurements can be used to project timber stands into the future. Questions remain, however, over using the data beyond its limits. Studies of this kind could be improved with additional and more consisted resource information.

INTRODUCTION

The aggregate timberland assessment system (ATLAS)[1] can simulate growth and harvest for a wide range timber types and management conditions. The ATLAS model was implemented to make 50-year timber supply projections by the USDA Forest Service in the 1989 resources planning act (RPA) timber assessment (U.S. Department of Agriculture 1989). The RPA objective is to compile a comprehensive assessment of the current and future National timber situation. This task required modeling growth and harvest on 344 million acres of privately owned commercial timberland. The ATLAS model was developed to replace TRAS (Larson and Goforth 1974), the timber model used for previous assessments. In the RPA modeling process, ATLAS was linked to the economic model TAMM (Adams and Haynes 1980). The resulting projections of timber inventories, harvests, stumpage prices, wood products consumption and production, and employment in the forest products sector were the result of interaction between biological and economic assumptions and the ensuing equilibrium between the timber supply and the demand for stumpage.

The timber side of this process begins with a periodic inventory of the private timberland in the Unites States conducted regionally by various Forest Inventory and Analysis (FIA) projects of the USDA Forest Service. The inventory system is designed to collect resource data from a state-level configuration of permanent plots on a 5- to 12-year cycle. To provide resource information for the this and subsequent Assessments, an RPA timber database was developed to includes 124,000 records representing approximately 483 million acres of available timberland (Waddell and Oswald 1986). This database proved crucial to the development of both the beginning inventory and projection parameters input into the ATLAS model. Database plot records included variables for state, ownership group, forest type, age class (for most plots), site productivity class, net growth, net volume, and plot expansion factors. Public land is included in the database, however, only data from plots on privately owned lands were included in the modeling process.

Three categories of input are required to make ATLAS timber projections. The first is a growing stock inventory that consists of acres and volume per acre by age class. The second is yield tables and growth parameters used in the projection mechanism. And third is harvest volume. The harvest request volume is determined exogenously, for RPA projections TAMM calculated a harvest request based on the demand for forest products. (A fourth category of input required for the RPA was projections of land area by forest type. Projections of forest type transition were supplied by Alig and others [U.S. Department of Agriculture 1989].)

---

[1] Mills, J.R. and Kincaid, J.C., [in preparation]: The aggregate timberland assessment system (ATLAS): a comprehensive timber projection model.

The options in ATLAS allow for harvested acres to get regenerated or leave the management unit. Area change options allow acres to change ownership, type, or management level. A detailed description of the harvest mechanism or area change options is beyond d the scope of this paper.

METHODS

For the purpose of the RPA timber projections, the continental U.S. was subdivided into 8 timber supply regions. Within each region, commercial timberland was identified by ownership, forest type, and site productivity class. These identifiers were the key variables used to sort and aggregate the growing stock inventory into management units. The management units represented 2 owner groups (forest industry and nonindustrial private), up to 9 forest types (per region), and 3 site productivity classes were carried for the Southeast, the South Central, and the Pacific Northwest Douglas-fir subregion. Within each management unit, growing stock is represented by an array of acres and volume per acre for 18 age classes (5-years per class in the South, 10-years elsewhere). Options allow for a stratification of the acres and volume within a management unit among 5 different management intensities. Each management intensity has growth parameters to model a response to a particular condition or treatment (custodial management, thinning, fertilization, genetic improvements, etc.). In the RPA timber projections, 249 management units were developed, 84 management units had more than one management intensity.

The model inputs for beginning inventories were either developed by the FIA units themselves, as was done in the North Central and South, or developed directly from the plot-level information in the RPA timber database. In either case, the starting area and volume was derived from FIA plot data.

Projecting Inventory Volume

The ATLAS projection methodology is built around the use of yield tables to project stand dynamics. Projections are for periods, the period length is consistent with the years in an age class. A stocking adjustment equation is the dynamic link between the existing timber inventory and the yield table. Stocking is growing stock volume divided by yield table volume. The stocking adjustment is based on the approach-to-normal concept, that is, over time stand density will naturally trend towards that of the normal undisturbed condition for that particular forest type and site (see Chaiken 1939, Gevorkiantz and Duerr 1938, Schumacher and Coile 1960). Growth is the volume change that results when acres are projected to the next oldest age class. A system similar to this was described by Chapman (1924) as a technique to employ a normal yield table to predict future volume for less than normal stands. Normal yield tables have typically not been used in ATLAS, it can be said the yield table acts as a volume "guide curve" rather than a volume asymptote. Instead of the term approach-to-normal, Leary and Smith (in press) use the term "relative density change" to describe this mechanism. The calculation of stocking within a single age class, its adjustment, and the subsequent change in volume can be represented as

$$S_{it} = V_{it}/Y_i$$
$$\text{when } Y_i > 0 \text{ and } i > 0 \text{ ,} \tag{1}$$

$$S_{i+1,t+1} = b_1 + b_2 S_{it} \text{ , or} \tag{2}$$

$$S_{i+1,t+1} = S_{it} + b_1 S_{it} + b_2 S^2_{it}$$
$$\text{when } b_1 > 0 \text{ and } b_2 < 0 \text{ , and} \tag{3}$$

$$V_{i+1,t+1} = S_{i+1,t+1} Y_{i+1} \text{ ;} \tag{4}$$

where
S = stocking ratio,
V = growing stock volume per acre,
Y = yield table volume per acre,
i = age class,
t = projection period, and
b_1, b_2 = stocking adjustment coefficients.

Equation (1) calculates the stocking ratio in the current period. Equations (2) and (3) are two forms of the equation used to calculate the approach-to-normal stocking adjustment. Equation (2) is a linear form derived from McArdle and others (1961). Equation (3) is a quadratic form derived by Gevorkiantz and Duerr (1938) and implemented in the North Central region by Leary and Smith (in press). Equation (4) represents the calculation of volume next period, acres associated with V have been assigned to the next highest age class. To define this single age class more specifically in an ATLAS inventory, subscripts could be added to represent management intensity, and management unit (b_1 and b_2 would not be subscripted by management intensity). Regenerated acres are assigned a beginning stocking ratio developed exogenously (discussed later).

The following formulation represents the summation of regular growth reported for period 't' within a single management intensity:

$$TRG_t = \sum_{i=0}^{n} V_{i+1,t+1} - V_{it} \text{ ;} \tag{5}$$

where TRG equals the total regular growth per acre calculated under one base yield table, n equals the highest age class. Total growth and total volume are the product of acres their and associated volumes per acre.

Developing Yield Tables

There were 422 yield tables used in the timber projections. They came from three sources: other models; earlier studies; and derived from plot data. As a model input, yield tables are a tabular representation of net volume per acre for 18 age classes. A difficulty encountered in the RPA was that published yield tables representing broad forest types rarely exist. Therefore, yield tables were developed to represent the various type-site-management aggregations. Most of the yield tables were not ownership specific, instead, growth differences between the two owners was reflected in the stocking adjustment mechanism and the regeneration stocking ratios.

Fully stocked yield tables were developed to represent the Southeast and South Central regions in the fourth forest study (U.S. Department of Agriculture 1988) (a predecessor of ATLAS; TRIM [Tedder and others 1987] was the timber projection model for the fourth forest study). These yields were derived from survey plot data that was screened to include only plots with enough growing stock to be classified as fully stocked by FIA standards (see McClure and Knight 1984). For these plots, the average net volume was calculated by type, site class, and age class (5-year). Curves were fitted to the averages for each type and site to create a smooth volume by age relationship. I would expect the yield volumes to range somewhere between classical empirical yield table volumes (an average volume calculated from all plots), and normal yield table volumes (calculated from undisturbed plots with normal stocking). Implementation of these yields for the RPA required recalibration of the stocking adjustment parameters.

The yields developed to project the Northeast, the Rocky Mountains, the Pacific Southwest, the Pacific Northwest Ponderosa Pine subregion,and several types in the Pacific Northwest Douglas-fir subregion were developed using growth data from the plots in the RPA timber database. In an earlier study I referred to yields derived from plot growth data as growth yields (Mills 1989). Growth yields were developed using all plots of the specific forest type (and site--when represented). An average growth per acre was calculated for each age class. Linear regression was used to smooth variation over the age classes and to predict growth for age classes containing no data. A yield table was developed by cumulating the average growth per acre over all 18 age classes. The value corresponding to each age class in the resulting table represents the summation of the average net growth up to and including the current age class.

The logic of growth yield tables is, at the regional level of aggregation a cross section of growth is an indicator of the average volume change and hence, makes a representative volume guide curve. The slope of fully stocked empirical yield curves tends to flatten in the older age classes, and this is often interpreted to indicate little or no net growth occurs in older stands. The plot growth associated with older stands, however, is often higher would be expected from the slope of fully stocked yields. The growth on older stands is built into the growth yield tables, making them compatible with the ATLAS "volume guide" role of yield tables.

The Calculation of Stocking

There are two types of inputs associated with stocking that were developed from the database. One is the parameters (b_1 and b_2) associated with equations 2 and 3, the other is a regeneration stocking ratio.

The parameters in equation 2 were calculated for all types and sites represented in the Northeast, Rocky Mountain, Pacific Southwest, and Ponderosa Pine Subregion of the Pacific Northwest. Elsewhere they were calculated selectively based on performance review of $b_1 = 0.1$ and $b_2 = 0.9$; these values asymptotically increment stocking to 1.00 . Equation 3 was used only in the North Central region.[2]

The parameters in equation 2 were estimated by owner and type (and site when enough plots existed). The procedure required the yield table, the plot volume per acre, and the annual growth associated with the plot (based on remeasurement). An estimate of plot volume was calculated for age = i-1 by subtracting net growth from net volume at age = i. A stocking ratio was then calculated for the plot volume at age = i (as in equation 1), and also for the estimate of volume at age = i-1 as shown below:

$$EV_{i-1} = V_i - G_i L \text{ , and} \qquad (6)$$

$$ES_{i-1} = EV_{i-1} / Y_{i-1} \text{ ;} \qquad (7)$$

where
i = age class
EV = estimated volume per acre
ES = estimated stocking
G = annual growth per acre, and
L = number of years per age class.

---

[2] For further information contact Rolfe Leary, North Central Forest Experiment Station, 1992 Folwell Ave., St. Paul, MN 55108

After screening the set of stocking ratios for outliers, liner regression was used to estimate b_1 and b_2 from the equation

$$S_i = b_1 + b_2 ES_{i-1} \, . \qquad (8)$$

Most regressions were run with at least 100 observations.

The regeneration stocking ratio is simply S_{it} for all t when i = 0, it represents the average stocking of the regenerated inventory associated with the yield table. For most yield tables the ratio was calculated from the inventory plot volume associated with the yield table. Average stocking was calculated for each age class and then acre weighted between age classes. This technique assumes regeneration will start at average stocking based on the inventory within the management unit (calculated when t = 0). After regeneration, stocking will trend towards the stocking asymptote. In the case when yield tables were developed to represent realized volume (i.e. under specific management regimes), the regeneration stocking ratios were 1.0 or 0.9 and the stocking asymptotes were 1.0 (the yield table).

The Assignment of Plot Age

In two regions the age variable was missing from approximately half of the plots. In both the Northeast and South Central regions, plots were not assigned a stand age because either the sampled stands were unevenaged, or the plot straddled two distinct forest conditions. In either case, plot age was not assigned because of differences in species, size, and age within the sample. Age is a required variable because it determines the initial starting point of acres in the volume projection mechanism. To estimate age for these plots, two age estimation techniques were employed. For the South Central region, researchers at the USDA Forest Service, Southern Forest Experiment Station (Starkville, MS) used yield tables and stocking variables to assign age based on plot volume and stocking. On Northeast plots that did have an assigned age, volume and age were poorly correlated. The assignment of age using yield table volume would seriously have truncated the distribution of acres by age class. To estimate the ages of these plots, I employed linear regression on variables such as stocking, site, average diameter, volume, and growth to developed estimation equations for plots with age (by forest type). These equations were then used to estimate the age of the indeterminant-age plots. Yield tables for the Northeast were developed from all plots after the assignment of age.

DISCUSSION

Many positive things should be said of overall RPA modeling effort. The communication with FIA units was good, the plot-level database was a new improvement, and the evaluation and comment received from reviewers was excellent. The extensive use of plot data for the development of growth and yield parameters made possible the Nationwide implementation of ATLAS. The RPA database was an essential tool; without which, given the timeframe and budgets, inputs for the Northeast and Rocky Mountain regions would not have been constructed. (An alternative for these two regions was to use the TRAS model with 10-year old input files.)

Data from permanent plots allows for a calculation and projection of change over time. If the latest measurements are considered the best evidence of what is happening in the field, then the use of growth data to calculate yield tables and density change parameters is consistent with our desire to project a future based on current observation. Not discussed here was the attempt to model assumptions of improvements in timber management. Several management intensities were developed to represent the expected returns from enrolling acres into intensive management schedules. The intensified management yield tables were either developed from models or based on other yield tables. The yield tables and acreage enrollment rates were subject to extensive review.

Several questions arose in the analysis and it is not clear how the answers would effect the results. Does a cross section of average growth really represent the potential for individual stands, or in this case, an aggregation of plots? If the values in a growth yield table are independent between age classes, then is this kind of yield table accurate for plot projections? The correlation of plot age with plot volume and growth might also be questioned. From the plot data there was no way to be sure how well the growth variable was correlated with the age variable. Disturbance or partial harvest of a plot could conceivably reduce the average age of a plot and the net plot growth might not be representative of the residual stand. A good correlation is important for the growth yield tables and stocking adjustment parameters. It is also not known what bias might have been introduced by the assignment of age to plots. Should the same technique be used to project both a true uneven age plot and a plot that straddles two different conditions? The answers to these questions are the subject of further study.

As good as the database was, there are some problems at the National level that are rooted in the FIA procedures themselves. Generally, there is a lack of consistency in the inventory techniques between FIA regions. The plot design, sampling intensity, inventory cycle, the calculation of growth, and the estimate of harvest vary by region. In many cases techniques have been tailored to particular local conditions and user requirements, in other cases the differences are based on tradition.

The Future

Preliminary work has already begun, the draft projections for the 1993 Assessment Update will be made in the summer of 1992. In the interim, there is time to make modest improvements. It has probably been said at the end of every study of this kind, a better job could be done with additional data. Any model used across such a large area will face many of the same data challenges. A list of desirable variables to incorporate in the next RPA database would include: (1) a disturbance code associated with the plot to indicate previous damage or harvest, (2) sample kind--to indicate a remeasured or newly established plot, and (3) age, volume, and growth estimated at time of previous survey. The historical information would allow for screening of plots that might exhibit growth that is not necessarily attributed to age or current volume. It would lead to finer calibration of the density change parameters, and in addition, provide data for the effort to improve the representation of harvest.

The assignment of age will be a recurring challenge for both uneven-aged stands and plots that straddle two distinct conditions. It has been proposed that an unevenaged projection system be developed for the Northeast. As the modeling process evolves to a finer level of resolution, a plot straddling two conditions will present a more difficult challenge.

If we keep in mind our purpose and our users, this modeling process will continue to evolve. New ideas translate into new models, additional validation and calibration, and there is always desire for "better" data. The projections for the 1989 timber assessment are finished, but the techniques and results of the whole process, from data collection through modeling, will be the topic lively discussion for years to come.

REFERENCES

Adams, Darius M.; Haynes, Richard W. 1980. The 1980 softwood timber assessment market model: structure, projections, and policy simulations. Forest Science Monograph 22. 64 p.

Chaiken, L.E. 1939. The approach of loblolly and Virginia pine stands toward normal stocking. Journal of Forestry. 37: 866-871.

Chapman, Herman Haupt. 1924. Forest mensuration. New York: John Wiley & Sons. 557 p.

Gevorkiantz, S.R.; Duerr, William A. 1938. Methods of predicting growth of forest stands in the forest survey of the Lake States. Economic Notes No. 9. St. Paul, MN: U.S. Department of Agriculture Forest Service, Lake States Forest Experiment Station. 59 p.

Knight, Herbert A.; McClure, Joe P. 1975. North Carolina's timber, 1974. Resour. Bull. SE-33. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 52 p.

Larson, R.W.; Goforth, M.H. 1974. TRAS: a timber volume projection model. Tech. Bull. 1508. Washington, DC: U.S. Department of Agriculture, Forest Service. 15 p.

Leary, Rolfe A.; Smith, W. Brad. [in press] Test of the TRIM inventory projection methods on Wisconsin jack pine. Can. J. For. Res. 20:

McArdle, Richard E.; Meyer, Walter H.; Bruce, Donald. 1961. The yield of Douglas-fir in the Pacific Northwest. Tech. Bul. No. 201. Washington, DC: U.S. Department of Agriculture, Forest Service. 74 p.

McClure, Joe P.; Knight, Herbert A. 1984. Empirical yields of timber and forest biomass in the Southeast. Res. Pap. SE-245. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 75 p.

Mills, John R. 1989. TRIM timber projections: an evaluation based on forest inventory measurements. Res. Pap. PNW-RP-408. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 27 p.

Schumacher, F.X.; Coile, T.S. 1960. Growth and yields of natural stands of the southern pines. Durham, NC: T.S. Coile, Inc. 115 p.

Tedder, P.L.; La Mont, Richard N.; Kincaid, Jonna C. 1987. The timber resource inventory model (TRIM): a projection model for timber supply and policy analysis. Gen. Tech. Rep. PNW-202. Portland, OR: U.S. Department of Agriculture Forest Service, Pacific Northwest Research Station. 82 p.

U.S. Department of Agriculture. 1989. An analysis of the timber situation in the United States: 1989--2040, parts I and II. Review Draft. Washington, DC: U.S. Department of Agriculture, Forest Service.

U.S. Department of Agriculture. 1988. The South's fourth forest: alternatives for the future. For. Resour. Rep. 24. Washington, DC: U.S. Department of Agriculture, Forest Service. 512 p.

Waddell, Karen; Oswald, Daniel. 1986. The National RPA timber database, 1990 RPA timber assessment data dictionary. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 65 p. On file with: Forest Inventory and Analysis Project, Forestry Sciences Laboratory, Pacific Northwest Research Station, P.O. Box 3890, Portland, OR 97208-3890.

Forest Change - An Assessment of Nationwide Forest Inventory Methods

Richard Birdsey
Forest Service - USDA
Washington, DC 20090-6090

Hans Schreuder
Forest Service - USDA
Fort Collins, CO 80526-2098

ABSTRACT

Forest Service Forest Inventory and Analysis (FIA) projects have developed several philosophies and techniques to estimate growth, removals, and mortality in different situations. Regional methods are related to past and current sampling designs and plot configurations, time between successive inventories, and regional characteristics of forest resources. Since FIA surveys involve an extensive network of permanent sample plots, some remeasured for more than 30 years, the data has unique value for long-term monitoring of forest conditions at the national level. Adoption of some new inventory techniques could improve regional and temporal consistency and enhance the usefulness of the data for monitoring purposes.

INTRODUCTION

This assessment is a summary of background material prepared for an internal review of Forest Inventory and Analysis (FIA) methods for assessing forest change. In the past, information from the nationwide forest survey was typically used to assess the adequacy of the nation's timber supply. Although timber supply is still an important issue, the public is demanding more information about current and prospective changes in forest resources. Increasingly, FIA is asked to provide non-traditional information such as the amount of carbon stored and accumulating in U.S. forests, or the effects of atmospheric pollutants on forest growth. In addition to traditional requests from forest industries and other timber interests for information about timber supplies, FIA has recently been getting requests for information related to environmental issues from the Environmental Protection Agency, the American Forestry Association, the President's Council of Economic Advisors, and the Man and the Biosphere Program, to name a few.

Our challenge is to make the best use of existing long-term data sets for answering new questions, and to develop better ways to meet these and future information needs. The basic inventory design was developed in the 1950's to provide efficient estimates of timber volume and forest condition. If current methods are inadequate or inefficient for meeting new information needs, then procedures will be revised in ways that maintain the strength of long-term continuity.

A BRIEF HISTORY OF FOREST INVENTORY METHODS

The systematic survey of U.S. forests was begun in 1930 under the authority of the McSweeney-McNary Forest Research Act of 1928. Sample surveys were made by what was known as the line-plot method (Hasel 1938; Robertson 1927). Parallel survey lines at fixed intervals were run at approximate right angles to the prevailing drainages. Quarter-acre sample plots were established at each 10-chain interval (660 feet) along the survey lines by field crews that literally walked from one end of a state to the other (Eldredge 1934). This design was used for most of the U.S. until about 1950, except during World War II when no surveys were conducted.

Of all the decades since the forest survey began, the 1950's had the most sweeping and lasting changes in inventory methods. During that decade, inventory projects in different regions began developing and applying new inventory designs based on advances in forest sampling theory and the widespread availability and use of aerial photography after the war (Hasel 1954). Cruise lines were largely abandoned in favor of permanent sample plots located in a systematic way. Information from aerial photos was used for stratification and plot location. The statistical technique of double sampling from aerial photographs was an important advance that is still widely applied today (Bickford 1952).

The plotless or variable-radius plot method of timber cruising, discovered by Walter Bitterlich, was introduced to the U.S. during the 1950's by Lew Grosenbaugh (Bitterlich 1948; Grosenbaugh 1952a, 1952b, 1958). Horizontal point sampling became generally accepted as a more efficient method for selecting sample trees for volume estimation. Inventory projects replaced fixed-radius sample plots with variable-radius sample plots. Various basal area factors and plot configurations were used. For example, in the South, pairs of 1/4 acre fixed-radius plots were initially replaced by pairs of variable-radius plots selected with a prism basal area factor of 10. But by the 1964 survey of Louisiana, the plot pairs established at each sample location were replaced by a cluster of 10 variable-radius plots selected with a prism basal area factor of 37.5, and supplemented with 3 small fixed-radius plots for sampling trees less than 5.0 inches in diameter at breast height (dbh). Similar changes were occurring at all the inventory projects across the country.

The establishment and remeasurement of permanent sample locations became popular during the 1950's under the label Continuous Forest Inventory (CFI). Permanent plots were seen as the answer to problems in estimating growth, removals, and mortality from temporary plots. Permanent plots were controversial, though, since higher establishment and marking costs made permanent plots more expensive than temporary plots for estimating all of the other information upon which decisions about actual forestry operations were based (Grosenbaugh 1959).

Despite the controversy, permanent, well-hidden sample plots became the norm across the U.S. Many of these same plot locations are still remeasured periodically, some for the fourth time. In the East, sample plots were located at the intersections of grid lines or at random locations selected from a systematic grid. In the West, mapping of forest areas and stratification from aerial photographs were used more frequently, often in combination with a double sampling design.

During the 1960's and 1970's, survey designs became increasingly diverse and complex. Sampling with partial replacement was developed as an efficient way to select an optimum mix of permanent and temporary plots (Ware and Cunia 1962). Advanced multi-stage and multi-phase sampling designs were developed to take advantage of new techniques in remote sensing. Different combinations of techniques were found suitable for each forest survey region. Many of these developments are evident in the survey designs currently used, but none has replaced double sampling from aerial photographs and permanent sample plots as the mainstay of the U.S. forest inventory design (U.S. Forest Service 1987).

OVERVIEW OF CURRENT FOREST INVENTORY METHODS

Inventory Objectives

Timber supply and environmental issues of the 1980's require good trend information and data for use in many kinds of models, in addition to the more traditional estimates of current conditions. Issues have increasingly involved detailed analyses of estimates from successive inventories. A recurring problem has been to determine whether observed trends are real or artifacts of changes in inventory and sampling design, and whether observed trends represent a departure from what is normal or expected. An elusive goal has been a credible statement about the causes of observed trends using periodic inventory data.

Sampling Design

Most FIA units use two-phase sampling for regular forest inventories. The first phase consists of a large number of sample points selected on aerial photographs and classified by a photo interpreter as, at minimum, forest and nonforest. The photo samples may also be classified by volume class, forest type, or disturbance. The second phase consists of smaller number of sample points selected on the aerial photograph and visited in the field. The field classification is typically used to correct for photo-plot misclassifications due to out-of-date photography or interpreter error. The sample points of phase two comprise the locations of permanent and temporary field plots used to estimate volume, growth, removals, mortality, and other forest inventory parameters, and to classify sample plots by ownership, forest type and other vegetation or site characteristics.

Typical variations among FIA projects involve the degree of dependence between phase one and phase two, the use of subsampling in either phase, the specific sampling rules used to select sample locations, and the sampling intensity. These variations affect the precision of estimates and the efficiency with which an inventory is conducted, but do not lead to any known bias in results.

Selection of Sample Trees

Sample trees are selected in clusters of fixed- and variable-radius sample plots at permanent and temporary locations. The vast majority of plot locations are now permanent. Generally, fixed-radius sample plots are used to select small trees (less than 5.0 in. dbh) or very large sample trees (greater than 35.0 in. dbh), while variable-radius plots are used to select other trees. From 4 to 10 sample points are arranged in a specified configuration to cover an area from 1 to 5 acres.

All FIA units rotate satellite sample points into a forest classification if the central or primary sample point is classified as forest. However, one significant difference among FIA units is the treatment of sample plots that straddle different forest conditions. Some projects move sample points within a cluster into the same forest condition, while other projects let the points stay where they originally fall.

Individual Tree Measurements

Individual tree measurements fall into 3 categories: those made for every sample tree, those made for a subsample of all sample trees, and those made for an independent sample of trees. Measurements made on subsamples or independent samples are usually used to develop models and estimate parameters that are applied to all sample trees. Examples are volume regressions, biomass regressions, diameter growth models, height equations, and utilization factors. In all cases the intent of the models is to relate something that is difficult or time-consuming to measure (such as volume) to variables that are relatively easy to measure over a large sample (such as dbh).

In general, every sample tree in the U.S. is identified for species and measured for dbh. Most have at least one length (or height) measurement. All sample trees are classified as growing stock or cull and assessed for damage. Depending on the region, additional measurements may be made, including section lengths, upper stem diameters, stump heights and diameters, cull deductions, and log or tree grade. Sometimes models are used to predict individual tree measurements that are missing or not made for every sample tree.

Standing or felled trees on subsample and independent sample plots may be measured in great detail and accuracy. Procedures and measurements vary by region and are usually adapted to local conditions and knowledge.

Tree Volume Estimation

Individual sample tree volumes may be estimated directly by detailed measurement of every sample tree, or by using a regression model with relatively simple and accurate tree measurements as independent variables. Estimates of volume growth are usually based on repeated volume estimates of sample trees surviving from one inventory to the next, or for new sample plots, estimates of volume increment derived from tree ring measurements or growth models. Estimates of mortality volumes are usually derived from past measurements of sample trees on remeasured plots and tree measurements at the second inventory that reflect tree dimensions at time of death. Estimates of removal volumes from sample plots are usually derived from past tree measurements on remeasured plots, and application of growth models to predict dbh and/or length at the time of removal.

Definitions and Estimators

National consistency is maintained through standard definitions of key terminology and variables. Certain estimation details may not be defined nationally. This is because specific inventory procedures are developed locally in order to make the best use of available sources of data, to be consistent with local data needs and standards, and to make the best use of local knowledge developed through many years of research.

Similarly, estimators are not prescribed nationally because they are always dependent on the sampling design, methods used to select sample trees, and availability of necessary models. In general the entire inventory system is adapted as well as possible to local conditions in order to achieve efficiency while serving local as well as national data needs.

The principal estimators of forest change are growth, removals, and mortality. These variables are defined as either average annual estimates for the intersurvey period, or current annual estimates. The second definition implies that successive inventory estimates are not linked, although permanent sample plots may still be remeasured to estimate growth, removals, and mortality. Current estimates are used when it is necessary to make all of the inventory estimates apply to a single point in time. Current estimates are generally preferred when the period between inventories is lengthy, when there have been significant changes in definitions or methods that make successive inventories incompatible, or when one of the estimates is derived from a different survey method such as a census of timber removals by forest industry. Estimates of periodic averages have become more common in recent years as the survey interval has shortened, improvements in data processing capabilities have allowed extensive recomputation of past surveys to account for procedural changes, and estimates of removals from sample plots rather than industry surveys have become more commonplace.

Classification Variables

All forest inventory projects collect and compile classification data used to post-stratify sample plots for analysis and screening. Examples of classifications are ownership, forest type, stand-size class, and stand age. Most classification variables have uniform national definitions, but definitions change over time, and the actual procedures to classify a sample plot may vary over time or from region to region.

Sample plot classifications are often derived from basic tree measurements, so the classification is dependent on the methods used to select sample trees. Sample plot design and rules for rotating sample points can therefore affect any plot classification, such as forest type or basal area, that is dependent on tree measurements. So changes in sample plot design and sampling rules can affect sample plot classifications in indirect ways.

USE OF FIA DATA FOR TREND ANALYSIS

FIA inventories represent one of the few sources of long-term measurement data, and the only comprehensive network of samples that truly represent forest conditions across most of the U.S. Other long-term data sets such as used in growth and yield studies typically cover selected forest types or management conditions, whereas FIA inventories cover all timberland regardless of location, condition, or management activity.

Despite the unique, comprehensive nature of the FIA sample network, it has been difficult for some researchers to accept FIA data as a source of valid trend information (Stout 1989). Part of the problem is a failure to correctly visualize inventory data as a repeated, natural experiment, and part is due to the lack of strict temporal consistency in inventory methods.

Diamond (1983) described 3 kinds of experimental approaches to ecological problems: laboratory, field, and natural experiments. In laboratory experiments the values of independent variables can be controlled, but ecological systems are usually so complex and variable over time that laboratory experiments are unrealistically simplistic. Field experiments can incorporate much of the natural variability of site and climate, but the experimenter can control only one or a few variables related to site or species, often at great cost. In a natural experiment, much control over independent variables is lost, leaving the experimenter with only site selection as a method of control. The goal of a natural experiment is to select sites as identical as possible in all factors except the one of interest. The major advantages of natural experiments are that a large amount of data can be gathered quickly and inexpensively, problems that cannot be created experimentally can be addressed, and problems involving large areas and the end results of long evolutionary processes can be studied. The major disadvantage is that with limited control of independent variables the experimenter is less sure that an outcome is the result of the specified factor or some unknown factor.

Most analyses of inventory results involve comparisons with the previous survey, and a conscientious analyst will diligently account for any changes in definitions or procedures. To fully analyze trends in growth or mortality, the analyst must reach back two inventories to establish initial conditions or screen data for consistent populations. Since 3 inventories typically span 20 or more years, the likelihood of changes in procedures and definitions affecting sample plot classifications is high. The analyst may need to spend considerable time and effort recomputing survey estimates from original data to effectively compensate for these changes.

Once a consistent set of data has been assembled, sample plots can be selected to control the effects of key variables for the two periods. In a cross-sectional study that related forest growth to sulfate deposition, Shifley (1987) selected sample locations from a large pool of inventory plots along a sulfate gradient and controlled the effects of growth variation due to stand age, site index, and stand density for 5 different forest types. This method of experimental control is also appropriate for temporal analysis of growth trends if the selection variables are strictly consistent.

CAN WE MEET FUTURE INFORMATION NEEDS?

For forest inventory, "acid rain", "global climate change", "forest decline" and "monitoring" are representative concerns of the 1980's. As foresters we are particularly interested in the future and how forests will change in response to long-term environmental changes, and what the potential of forests may be to mitigate the effects of anticipated changes.

To address these issues will require an ability to separate expected forest changes due to natural processes, timber harvest, and management, from unexpected changes due to anthropogenic environmental changes. We can anticipate a future need to look back at today's data to establish baseline information for evaluation of future forest conditions.

The national forest inventory needs a strengthened core of basic data elements that are accurately measured in a consistent way to effectively analyze future forest changes. Sample plot classifications should be derived from basic measurements that are repeatable over time by different field crews. The sampling methods must also have temporal consistency because a plot classification may change under different sampling rules. We must be careful about making changes to achieve spatial consistency that cause basic data to be incompatible with the past or the future.

So when we think about changing sample plot designs or methods of classifying sample plots, we should always consider how this will affect the usefulness of the data 20 or 60 years from now. The inventory methods developed in the 1950's, and still in place today as the core of the national inventory, represent our basis for analyzing long-term forest changes. Will an analyst from 2020 curse us because we changed something to be more efficient or use a more up-to-date technique? If we are right in anticipating that long-term environmental considerations are of immediate and future concern, then it might be wise to remember temporal consistency when contemplating a state-of-the-art change in inventory procedures.

REFERENCES

Bickford, C.A., 1952: The sampling design used in the forest survey of the Northeast. J. For. 50: 290-293.

Bitterlich, W., 1948: Die winkelzahlprobe. Allgemeine Forst-und Holzwirtschaftliche Zeitung 59(1/2): 4-5.

Diamond, J.M., 1983: Laboratory, field, and natural experiments. Nature 304: 586-587.

Eldredge, I.F., 1934: The forest survey in the South. Forest Survey Release No. 1. New Orleans, LA: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station, 5 p.

Grosenbaugh, L.R., 1952a: Plotless timber estimates -- new, fast, easy. J. For. 50: 32-37.

Grosenbaugh, L.R., 1952b: Shortcuts for cruisers and scalers. Occasional Paper 126. New Orleans, LA: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station, 24 p.

Grosenbaugh, L.R., 1958: Point-sampling and line-sampling: probability theory, geometric implications, synthesis. Occasional paper 160. New Orleans, LA: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station, 34 p.

Grosenbaugh, L.R., 1959: Should continuity dominate forest inventories? Paper presented at short course on Continuous Inventory Control in Forest Management, U. of Georgia, May 4-8. 13p.

Hasel, A.A., 1938: Sampling error in timber surveys. J. of Agric. Res. 57(10): 713-736.

Hasel, Austin A., 1954: Problems in forest inventory: from the statistical point of view. In Statistics and Mathematics in Biology, Ed. by Kempthorne et. al. Ames, Iowa: The Iowa State College Press. 263-272.

Robertson, W.M., 1927: The line-plot system, its use and application. J. For. 25: 157-163.

Shifley, S.R., 1987: Analysis and modelling of forest growth trends along a sulfate deposition gradient in the north central United States. In proceedings, Forest Growth Modelling and Prediction, Ed. by Ek et. al., Minneapolis, MN: Society of American Foresters, 506-513.

Stout, B.B., 1989: Forest decline and acidic deposition - a commentary. Ecology 70(1): 11-14.

U.S. Forest Service, 1987: Forest Service Resource Inventory: An Overview. Washington, DC: U.S. Department of Agriculture, Forest Service, 29 p.

Ware, K.D. and Cunia, T., 1962: Continuous forest inventory with partial replacement of samples. For. Sci. Monograph 3, 40 p.

Fixed-Versus Variable-Radius Plots for Change Estimation

Charles T. Scott, Research Forester
James Alegria, Forester
USDA Forest Service
Northeastern Forest Experiment Station
5 Radnor Corporate Center, Suite 200
Radnor, PA 19087

ABSTRACT

Much forestry literature has focused on fixed-area versus variable-radius plots for estimation of current forest characteristics. Variable-radius samples were found to be most efficient for current basal area and volume estimates, but much controversy arose over their efficiency for change estimation. As an empirical test of the cost effectiveness of each method, 14.8-ft plots were compared with 26.5-ft^2/ac BAF (basal area factor) plots. Each pair was remeasured after 12–13 years and detailed cost measurements recorded for all aspects of the survey. With the exception of removals, fixed-area plots were more cost effective than variable-radius plots for all attributes involving number of trees. Fixed-area plots also were more efficient for basal area and volume estimation of ingrowth and mortality, and the estimation of net change and net growth of basal area. Variable-radius plots were more efficient for estimating all attributes of net cubic-foot volume, except ingrowth and mortality. Variable-radius plots also were more efficient for estimating the number of trees and basal area of removals and for current basal area.

INTRODUCTION

Shortly after Grosenbaugh (1952) introduced variable-radius (Bitterlich point) samples to forestry in the U.S., a number of investigators began comparing this system to fixed-area (circular plot) sampling. Examples are Grosenbaugh and Stover (1957), Arvanitis and O'Regan (1967), Martin (1983), Oderwald (1981), and Ranneby (1980). Under most conditions, point samples were found to be more cost effective than fixed plots for estimates of current basal area and volume. Fixed plots were usually found to be superior for estimates of frequency (trees/acre). Oderwald, studying basal area estimates using a mathematical approach, contended that fixed plots were better with a square-lattice spatial distribution and point samples were better with random or clustered distributions. Of these authors, Arvanitis and O'Regan, Ranneby, and Martin compared the two inventory methods for change estimation.

Martin used a mathematical approach to investigate forest growth attributes. He concluded that if tree volume growth is negatively correlated with basal area, fixed plots are always best. With positively correlated growth variables, the coefficient of correlation must be greater than some critical value for point samples to outperform fixed plots. The critical value is a function of the diameter and spatial distribution of the trees.

Barrett and Carter (1968), Ek et al. (1984), Kulow (1966), Mesavage and Grosenbaugh (1956), Nyyssonen and Kilkki (1965), Scott (1981), Scott et al. (1983), and Shirley (1960) used cost data when determining the most efficient method. Precise methods may not result in efficient methods. Comparisons must be made on either precision when cost is fixed, or on cost when precision is fixed. The objective of this paper is to compare fixed-plots and point samples on a fixed-cost basis for a variety of attributes, but focuses primarily on components of change.

METHODS

Twenty dependent variables (attributes of interest) were identified. They are the current values of, or the changes in, the number of trees, basal area, and net cubic foot volume per acre. The changes are further classified by ingrowth, accretion, mortality, removals, net growth, and net change. Ingrowth represents the current value of trees which became merchantable ($\geq$ 5" diameter) during the period and accretion the value at time 2 minus the value at time 1 using the current number of trees per acre for both survivor and nongrowth trees (Roesch et al. 1989). Nongrowth trees were merchantable at time 1, but were not measured until time 2. Time 1 values were determined using diameter as estimated (predicted) from a diameter growth model (Roesch et al. 1989). Mortality and removals are defined using the value at time 1. Net growth represents ingrowth plus accretion minus mortality. Net change is net growth minus removals.

In another study, data were collected in Hancock County, Maine, during 1968–69 and in 1981 on fifty-eight co-located 1/5-acre and 10-point clusters (37.5 ft^2/ac BAF) with the 1/5-acre fixed plot being superimposed over the first point (fig. 1). The 1/5-acre plots were composed of two concentric circular plots: 1/10-acre for poletimber and 1/5-acre for sawtimber. In 1981 observation time for each item tallied was measured by a third crew member for both plot types. Distance, azimuth, species, diameter at breast height (dbh), merchantable height, percent cull, tree history, and some attributes not used in this study were recorded for all trees. Because distance and azimuth to each tree were recorded, we were able to test other plot sizes and BAF's.

Fixed Cost Comparison

Because only one fixed plot was observed at each location, comparisons were made only for a single fixed plot versus a single co-located point sample, rather than the cluster of points. Using a 37.5 BAF results in slightly more than two trees per point—not a very strong basis for comparison. Since distance and azimuth data were available, we used the concentric 1/5-acre-plot data to lower the effective BAF, increasing the number of trees sampled. The smallest BAF which ensured that the largest tree in the population (31.2 in dbh) would fall within the 1/5-acre plot was 26.5. The 26.5 basal area factor was large enough that the limiting distances for all poletimber-size trees were within the 1/10-acre plot. The use of these plots sizes is not a recommendation—the plots were chosen to provide direct comparison using this data set. Other comparison methods could have been used.

Cost Models

From the time data, cost models were developed as functions of plot size, mix of poletimber and sawtimber trees, computed walking distance between trees, and initial fixed time to begin measuring the first tree on the plot. The number of trees within a plot increased proportionately with plot size. Average tree measurement time were 2.11 minutes for poletimber and 3.56 minutes for sawtimber. The additional time required for measurement of a sawtimber tree is due to log grade and sawlog height which are not collected for poletimber. The walking distance between trees changed linearly with the square root of plot size. The point-sampled trees also were more

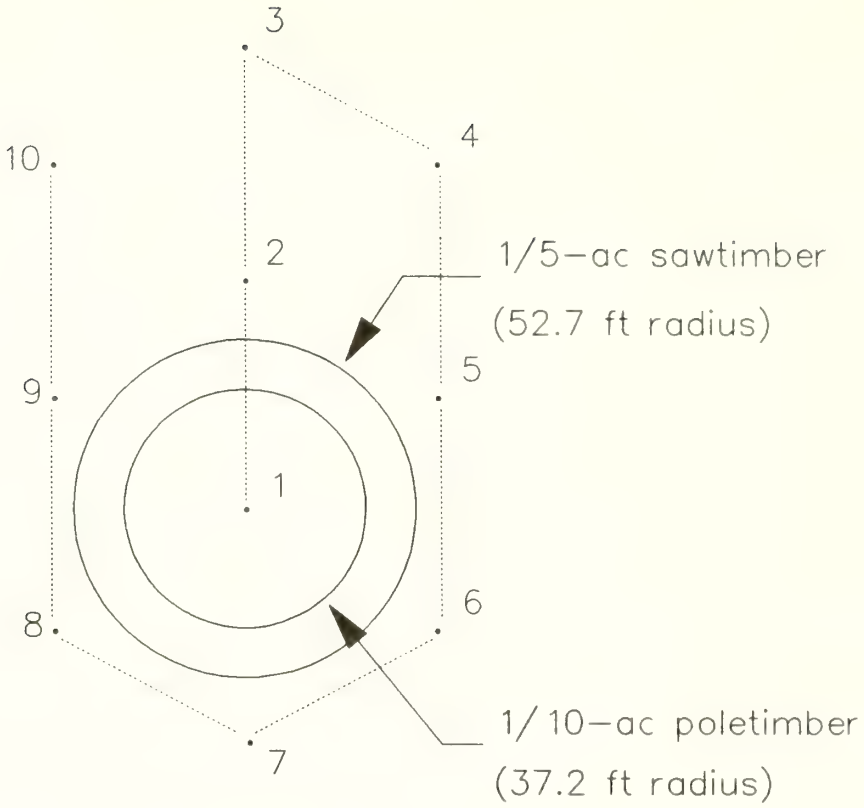


Figure 1. Layout of poletimber and sawtimber fixed plots co-located with cluster of 37.5 BAF point samples spaced 70 feet apart.

scattered than those sampled with fixed plots. The time model (in minutes) to walk from one tree to the next was:

$$T_f = 0.75 + 2.11\ N_{pole} + 3.56\ N_{saw} + N\ (0.521 + 1.29\ z^{0.5}) \qquad (1)$$

for measuring fixed plots of size, z (in acres), and:

$$T_p = 0.99 + 2.11\ N_{pole} + 3.56\ N_{saw} + N\ (0.170 + 2.22/BAF^{0.5}) \qquad (2)$$

for measuring point samples. N_{pole} and N_{saw} are the number of poletimber and sawtimber trees per point, respectively. N is their sum. The fixed cost of starting to measure trees on a plot was 0.75 minutes for fixed plots and 0.99 minutes for point samples.

The time expected to measure a 26.5-BAF point was estimated, and a fixed plot size determined that would require the same measurement time. Applying the 26.5-BAF prism to the data, an average of 2.56 poletimber and 1.82 sawtimber trees were measured per point for a total of 4.38 trees. Using the cost data, the predicted time was 15.50 minutes. By iteratively adjusting the fixed-plot radius and applying the fixed-plot cost data, a radius of 14.8 ft was found to yield a time of 15.46 minutes. This translates to a fixed plot of 0.0158 ac. An average of 3.98 poletimber and 0.845 sawtimber trees were measured per 14.8-ft plot, for a total of 4.825 trees. The average distance between trees for 26.5-BAF points was 12.7 ft versus 9.3 ft for the 14.8-ft plot.

The criterion for judging the plot designs for each of the 20 dependent variables was relative efficiency. Relative efficiency of point sampling is defined as:

Relative efficiency = fixed-plot variance/point-sample variance.

A relative efficiency of less than one favors the fixed plot and greater than one, the point sample.

RESULTS

Fewer trees were sampled on the 26.5-BAF plot (254) than on the 14.8-ft plot (278). The point samples have a greater proportion of the more time-consuming sawtimber trees than the fixed plots. The point-sampled trees also are more scattered than those sampled with fixed plots, thus tree location times are greater. For point sampling to be more efficient than fixed sampling, the reduction in variance due to the probability proportional to size estimation must be greater than the increase in variance due to sampling fewer trees.

The means, variances, and relative efficiency are given in Table 1. The means indicate close agreement between the two methods, none of the differences approaches statistical significance. Only the means of net growth and net change do not appear to match well due to the negative components in the estimator (mortality and removals).

On the basis of the relative efficiency, all attributes except removals, favored fixed plots for the number of trees per acre. Ingrowth and mortality were more efficiently estimated with fixed plots, even for basal area and volume. Ingrowth and most mortality occurred in the smaller diameter classes (see fig. 2), thus were better estimated with fixed plots which sample a higher proportion of small trees than do point samples. The strong influence of the ingrowth and mortality components resulted in better estimates for basal area of net growth and net change with fixed plots.

Removals were better estimated with point samples; most of the removals were in the larger diameter classes, thus point samples selected a greater proportion of them than did fixed plots, thereby increasing the relative efficiency. Most of the attributes involving volume were more efficiently estimated with point sampling. Only for ingrowth and mortality volumes did fixed plots dominate. Point samples had a smaller variance for basal area estimation of accretion, removals, and current values.

CONCLUSIONS AND SUMMARY

The results confirm some past findings. Current number of trees is best estimated with fixed plots. Current basal area and volume are best estimated with point samples.

Fixed plots were more efficient for:

Table 1. Relative efficiency (Rel. eff.) of 26.5-BAF point samples to 14.8-ft radius fixed plots by attribute (cost fixed).

	Variance		Rel.	Mean	
Attribute	Point	Fixed	Eff.	Point	Fixed
		Trees/Acre			
Ingrowth	12515	7575	0.60	64.87	78.00
Mortality	12263	5598	0.46	52.69	50.87
Removals	558	645	1.16	7.73	9.04
Net growth	26864	15424	0.57	12.17	27.13
Net change	27543	15549	0.56	4.44	18.09
Current	25893	16974	0.66	258.16	254.34
		Basal area/Acre			
Ingrowth	345	230	0.67	10.96	13.02
Accretion	228	291	1.28	22.02	21.64
Mortality	596	380	0.64	13.70	12.94
Removals	200	303	1.52	4.57	5.27
Net growth	1066	855	0.80	19.28	21.72
Net change	1315	1004	0.76	14.71	16.45
Current	2727	3391	1.24	96.38	91.86
		Cubic foot volume/Acre			
Ingrowth	43696	38871	0.89	125.56	155.43
Accretion	203084	277200	1.36	376.75	408.41
Mortality	128022	81480	0.64	195.88	176.34
Removals	68928	91220	1.32	78.03	84.44
Net growth	347777	429631	1.24	306.44	387.50
Net change	440085	457673	1.04	228.40	303.06
Current	1167549	1724301	1.48	1634.34	1593.73

1. Number of trees per acre (ingrowth, mortality, net growth, net change, and current values—except removals),
2. Ingrowth (number of trees, basal area, and volume),
3. Mortality (number of trees, basal area, and volume),
4. Net growth and net change in basal area.

Point samples were most efficient for:

1. Net cubic-foot volume per acre (accretion, removals, net growth, net change, and current values—except ingrowth and mortality),
2. Removals (number of trees, basal area, and volume),
3. Current basal area.

Of the 20 attributes investigated, 11 were more efficiently estimated by fixed plots and 9 by the point samples. However, most of the volume attributes were better estimated with point samples.

The results shown are for a single population in Maine using one estimation method (Roesch et al. 1989) for point samples. Only one co-located point with a single plot size for each plot type was investigated. Further comparisons should be made for other plot sizes, for clustering

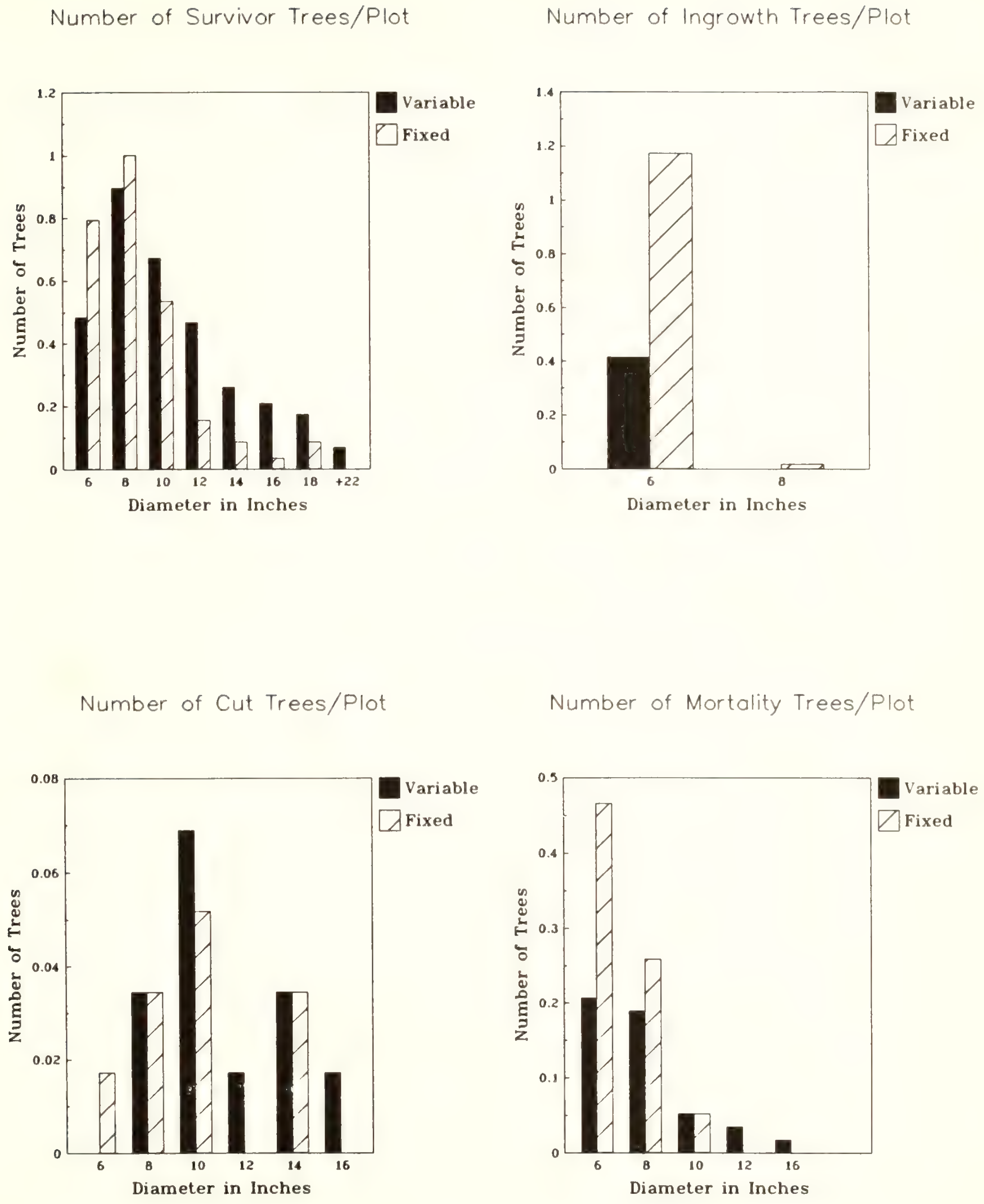


Figure 2. Number of trees sampled on 26.5-BAF point samples and 14.8-ft radius fixed plots by diameter class for survivor, ingrowth, cut, and mortality.

of plots, and for other diameter and spatial distributions. However, the patterns probably would be repeated in other studies.

Attributes concentrated in the small diameter classes, or are not correlated with current basal area probably would be more efficiently estimated with fixed plots. While both mortality and ingrowth were concentrated in the small-diameter classes in this study, ingrowth could be shifted to large-diameter classes when time between measurements is long or growth rates are high. Also, if most stands are mature, then mortality would be shifted to the large-diameter classes. Attributes concentrated in the large-diameter classes, or correlated with current basal area, are likely to be most efficiently estimated with point samples.

REFERENCES

Arvanitis, L.G., O'Regan, W.G., 1967: Computer simulation and economic efficiency in forest sampling. Hilgardia, vol. 38, pp. 133–164.

Barrett, J.P. and Carter, J.F., 1968: Comparing methods of sampling a white pine stand. Journal of Forestry, vol. 66, no. 11, pp. 861–863.

Ek, A.R., Scott, C.T., Zeisler, T.R., and Benessallah, D., 1984: Cluster sampling: from theory to practice. In: Inventorying forest and other vegetation of the high latitude and high altitude regions: Proceedings of the conference; Fairbanks, Alaska, Society of American Foresters, 84-11. pp. 78–81.

Grosenbaugh, L.R., 1952: Plotless timber estimates—new, fast, easy. Journal of Forestry, vol. 50, no. 1, pp. 32–37.

Grosenbaugh, L.R. and Stover, W.F., 1957: Point sampling compared with plot sampling in southeast Texas. Forest Science, vol. 3, no. 1, pp. 2–14.

Kulow, D.L., 1966: Comparison of forest sampling designs. Journal of Forestry, vol. 64, no. 7, pp. 469–474.

Martin, G.L., 1983: The relative efficiency of some forest growth estimators. Biometrics, vol. 39, no. 3, pp. 639–650.

Mesavage, C. and Grosenbaugh, L.R., 1956: Efficiency of several cruising designs on small tracts in north Arkansas. Journal of Forestry, vol. 54, no. 9, pp. 569–576.

Nyyssonen, A. and Kilkki, P., 1965: Sampling a stand in forest survey. Acta Forestalia Fennica, vol. 79, no. 4, 20 pp.

Oderwald, R.G., 1981: Comparison of point and plot sampling basal area estimators. Forest Science, vol. 27, no. 1, pp. 42–48.

Ranneby, B., 1980: The precision of the estimates of the growth in basal area. Swedish University of Agricultural Sciences, Section of Forest Biometry, report no. 19, 25 pp.

Roesch, F.A., Jr., Green, E.J. and Scott, C.T., 1989: New compatible estimators for survivor growth and ingrowth from remeasured horizontal point samples. Forest Science, vol. 35, no. 2, pp. 281–293.

Scott, C.T., 1981: Design of optimal two-stage multiresource surveys. Ph.D. dissertation, University of Minnesota, St. Paul, 138 pp.

Scott, C.T., Ek, A.R., and Zeisler, T.R., 1983: Optimal spacing of plots comprising cluster in extensive forest inventories. Proceedings of Renewable Resource Inventories for Monitoring Changes and Trends, College of Forestry, Oregon State University, Corvallis, Society of American Foresters, 83-14, pp. 707–710.

Shirley, F.C., 1960: A comparison of the amount of time needed to measure fixed and variable radius cruise plots. M.S. thesis, College of Forestry, State University of New York, Syracuse, 61 pp.

Assessing Pine Regeneration for the South Central United States

William H. McWilliams
Research Forester

USDA, Forest Service
Southern Forest Experiment Station
Box 906, Starkville, MS 39759

ABSTRACT

Poor regeneration of pine following harvest on nonindustrial timberland has been identified as a major cause for loss of pine forests and slowdown of softwood growth in the Southern United States. Developing a strategy for regeneration assessment requires clear definition of sampling objectives, sampling design, and analytical processes. It is important that regeneration surveys include enough flexibility to assess the diversity of cover types and management intensities common on upland pine sites. Discussion centers on the USDA, Forest Service's successive forest inventories in the South Central Region.

INTRODUCTION

Pine regeneration has become a dominant issue confronting the forestry community in the Southern United States because of a thirty percent decrease in the acreage of nonindustrial pine stands since the early 1950's. The latest inventory results show a resultant slowdown in aggregate softwood growth, which is unusual in the South. Poor regeneration of pine following harvest on nonindustrial timberland has been cited as a major cause of this decline (USDA, Forest Service 1988).

Although the importance of monitoring pine regeneration is evident, there are few guidelines for implementing large scale regeneration surveys. However, a few reports have proven useful for developing a strategy (Stein 1984a,b; Kaltenberg 1984). The purpose of this paper is to document procedures used by the USDA, Forest Service, Southern Forest Experiment Station, Forest Inventory and Analysis unit (SO-FIA). The SO-FIA conducts successive forest inventories in Alabama, Arkansas, Louisiana, Mississippi, Oklahoma, Tennessee, and Texas. Discussion centers on sampling objectives, sampling design, stocking assessment, and analytical techniques for estimating the status of regeneration following harvest. Results of regeneration surveys are reported elsewhere (McWilliams 1989).

SAMPLING OBJECTIVES

Regeneration surveys typically supply input for yield models or measure the success of silvicultural treatments for a particular class of forest. In contrast, the main goal of the SO-FIA inventories is to quantify conditions found over a wide range of forest types, productivity classes, stocking levels, and management intensities. Most users of regional regeneration statistics are analysts and policymakers interested in the impact of reforestation incentive programs or predicting future timber supplies.

Specific objectives of the SO-FIA regeneration survey have been:

-To determine the broad status of regeneration efforts and allow comparisons of success among forest types, productivity classes, and ownerships.

-To identify long term trends that reflect the combined effects of weather, insects, disease, fire and other factors.

-To predict the future species composition of harvested stands.

Additionally, the methodology of the regeneration sample was intended to:

-Achieve consistency with other measurements and definitions.

-Provide enough flexibility to assess planted, seeded, and residual trees.

-Maximize utility of the data and minimize the costs of collection.

SAMPLING DESIGN AND STOCKING

The SO-FIA currently uses a two-phase sample of temporary aerial photo points and a systematic grid of permanent ground plots. The area of forested land is determined by photointerpretation of temporary points and field checks of permanent plots. Field measurements are conducted on a subset of permanent plots spaced 3 miles apart. Tree data are collected on measurement plots that are currently forested or were forested at the time of the previous inventory. Tree data are used to estimate volumes, basal area, numbers of trees, and stocking per acre. Per acre estimates are expanded using factors derived as part of the forest area determination.

Each measurement plot is comprised of a cluster of 10 satellite points spread over an acre of forest (figure 1). At each point, trees 5.0 inches in diameter at breast height and larger are selected for measurement on a variable radius plot defined by a 37.5 BAF prism. Trees from 1.0 to 4.9 inches in diameter are tallied on 1/275 acre fixed plots at the first three points and at any remaining points where fewer than two trees 5.0 inches in diameter or larger are tallied. If no trees greater than 1.0 inch are tallied at a point, then seedlings are tallied.

Stocking percentage is computed by comparing the tree tally to a standard for fully stocked forests:

Diameter (in inches)	No. of trees per acre	Diameter (in inches)	No. of trees per acre
seedlings	600	16	72
2	560	18	60
4	460	20	51
6	340	22	42
8	240	24	36
10	155	26	31
12	115	28	27
14	90	30 +	24

All live trees are assigned a stocking percent in relation to the standard. At each of the 10 points, stocking percent is summed for the largest and most dominant trees until 16 percent is reached (based on a biological maximum of 160 divided by 10 points). This distributes stocking estimation

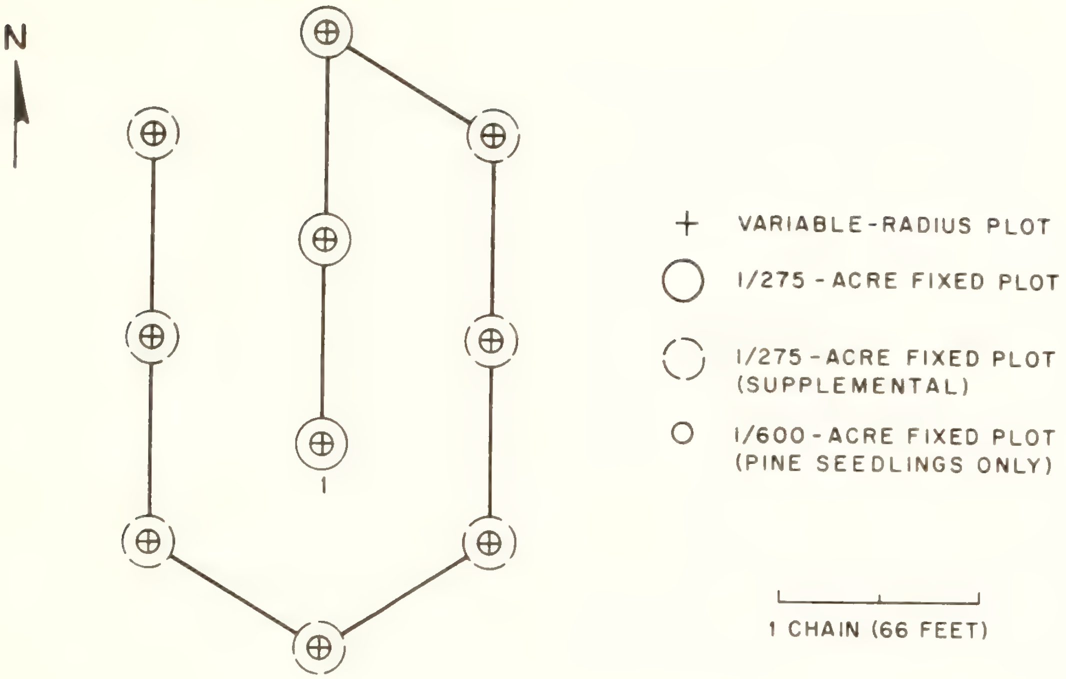

N
1
+ VARIABLE-RADIUS PLOT
1/275-ACRE FIXED PLOT
1/275-ACRE FIXED PLOT (SUPPLEMENTAL)
1/600-ACRE FIXED PLOT (PINE SEEDLINGS ONLY)
1 CHAIN (66 FEET)

Figure 1. Sample plot configuration.

across the sampled acre. Summation of stocking for all points produces SO-FIA's estimate of stocking for the plot. For each plot, the stocking plurality by species and dominance class forms the basis for forest type classification.

The SO-FIA measurement plots were not designed to sample pine seedlings, per se. One common perplexity is for dominance to shift from hardwoods to pine during the early years of stand development. To account for this phenomenon, a 1/600 acre fixed plot is established at each point. Each plot is checked for the presence of at least one well established, free-to-grow pine seedling. For each plot with a pine seedling, 10 percent is added to the stocking of pine that is used in the pine regeneration assessment.

PINE REGENERATION ASSESSMENT

Establishing a useful technique for measuring pine regeneration requires that some basic decisions be made regarding: 1)What timberland needs assessment? 2)What trees should be included as part of the regenerated stand? and 3)What constitutes an adequate level of stocking? Answers to these questions are complicated by the large scale of the inventory, but as Stein (1984a) has indicated, adequate definition of sampling objectives clarifies most of the issues.

Regeneration assessments are usually limited to sites with proven capability for growing pines, referred to as upland pine sites. Stand disturbance is classified to determine the area in need of assessment. The SO-FIA uses three disturbance categories: commercial harvest, intermediate stand management, and natural disturbance. Classification of commercial harvest is based on cutting intensity and inspection of the residual stand:

Clearcut - removal of all growing stock and/or rough and rotten trees.

Partial cut - pine selection cuts, diameter limit cuts, highgrading, or any other practice that tends to remove the most marketable trees, leaving a residual stand of growing stock and/or rough and rotten trees.

Other harvest - seed tree cuts, shelterwood cuts, and salvage cuts.

Pine regeneration assessment focuses on clearcut and partial cut timberland. Other harvests can be examined after residual trees are felled. Timberland that was cleared for nonforestry use is excluded.

The SO-FIA includes trees of all size classes in the regeneration assessment based on management practices of the region, the "snapshot" nature of the sample, and the need for consistency with other SO-FIA estimates. The most prevalent harvest practice in the South Central region is partial cutting, primarily due to its use by nonindustrial private owners who rely heavily on natural regeneration (McWilliams 1989). Thus, pole and sawtimber size trees need to be included for consistency with SO-FIA's concept of area occupancy, which underlies all stocking related variables.

Residual pine stocking is evaluated for the growing stock component of the post harvest stand, whether the trees originated from planting, seeding, or were left from the previous stand. This deviates slightly from forest type and stand size variables that are based on all live trees. The growing stock limitation ensures that only trees with management potential are included. The difference between using all live and growing stock trees is minor in terms of the results.

Limiting the regeneration assessment to nonsuppressed trees helps to improve the ability to predict future species composition. As indicated in the previous section, SO-FIA's stocking calculation is oriented towards dominant trees of the existing canopy. Overtopped trees can be screened out of the stocking summation.

DATA ANALYSIS

Pine stocking is grouped into classes to provide a relative measure of success:

Low - less than 30 percent stocked with pine.

Medium - from 30 to 59 percent stocked with pine.

High - at least 60 percent stocked with pine.

Pine stocking classes furnish the analyst with an indication of future species composition. A stand with a high stocking of pine will usually develop into a pure pine stand. Some medium stocked stands will develop into pure pine stands, while others will contain a mixture of pine and hardwoods. In the absence of later hardwood control, stands with a low stocking of pine usually evolve into hardwood stands. Pine stocking classes provide more information than the existing forest type classifications because of shifts in species plurality that often occur soon after harvest.

For analytical purposes, pine stocking classes are used in conjunction with other plot level descriptors. Regeneration success, on a regional scale, is related to ownership and preharvest forest

Table 1.--Area of harvested pine-site timberland retained as forest by ownership, forest type, and pine stocking class.

Ownership and forest type	Past area	Total	Pine stocking class Low	Medium	High
	--------------------Million acres-----------------				
Industrial					
Pine	9.4	4.4	0.7	1.0	2.7
Mixed	4.4	2.0	0.5	0.6	0.9
Hardwood	3.3	1.4	0.6	0.3	0.5
Total	17.1	7.8	1.8	1.9	4.1
Nonindustrial					
Pine	14.9	4.9	1.4	1.2	2.3
Mixed	10.0	2.9	1.3	1.0	0.6
Hardwood	12.4	2.7	0.9	0.5	0.3
Total	37.3	10.5	4.6	2.7	3.2
All private					
Pine	24.3	9.3	2.0	2.2	5.0
Mixed	14.4	4.9	1.8	1.6	1.5
Hardwood	15.7	4.1	2.5	0.8	0.8
Total	54.4	18.3	6.4	4.6	7.3

Source: The most recent forest inventories of Alabama (1982), Arkansas (1988), Louisiana (1984), Mississippi (1987), Oklahoma (1986), and Texas (1986). Industrial timberland includes land under long-term lease from nonindustrial private owners.

type. In general, pine regeneration is more successful on industrial timberland and following harvest of pine forest types.

To illustrate, table 1 shows the distribution of harvested timberland by ownership, preharvest forest type, and pine stocking class based on a study conducted in six states (McWilliams 1989). The results indicate that 61 percent of the harvested pine type timberland owned by forest industry had a high stocking of pine. The estimate of success increases to 84 percent if stands with medium stocking are included. Adding the classes together may be justified because of the intensity of harvesting on industrial timberland. Nearly half of the pine type timberland was harvested (compared to roughly one-third of the nonindustrial pine forest).

For nonindustrial private owners, success estimates for harvested stands ranged from 47 percent for high stocking to 71 percent for high and medium stocking combined. Including medium stocked stands may also be logical for nonindustrial owners because they often have limited investment capital, and mixed stand management offers a viable low cost alternative (Cain 1988, Phillips and Abercrombie 1987). Although a relative measure of success is often preferred, some applications require a single statement regarding regeneration success. In such a case, it is useful to refer to the overall replacement rate for harvested pine stands. The replacement rate is computed by dividing the area with a high stocking of pine by the area of pine stands that was harvested. In the example, the rate is 93 percent for forest industry and 65 percent for nonindustrial owners.

Percentages must be used with caution for trend analysis. For example, a static replacement rate between two inventories could imply no change in the total area regenerated. This is not the case if the area harvested increased during the inventory period.

The SO-FIA's systematic sampling design provides a spatial dimension to the analyses. Figure 2 depicts a sequence of maps showing the regeneration situation for nonindustrial owners in Mississippi. Screening the sample of pine site plots for harvesting, residual pine stocking, and site productivity highlights the regional dimension. If the most productive sites are a priority, the southern half of the State has the largest area in need of pine regeneration.

DISCUSSION

Successive forest inventories reflect the period prior to measurement (inventory periods are usually 7 to 10 years). As such, regeneration surveys measure the effects of economic cycles, harvest intensity, weather, disease, and other factors. This is an important advantage in contrast to the limitations associated with using planting records to gauge regeneration following harvest. Planting records do not reflect plantation failure and exclude natural regeneration. Drought and fusiform rust infection have been significant factors in recent years (Burnett 1987, Lenhart et al. 1988). An informal study in Texas predicted a 60 percent loss of seedlings planted in 1988 due to the severe summer drought (Cohen 1988). Also, planting records include significant areas planted on marginal cropland as part of the Conservation Reserve Program. Such plantings do not contribute to regeneration of harvested stands.

Users of regeneration survey results need to be cognizant of the implications of estimation procedures. Applying a single stocking standard to a range of management intensities and site conditions imposes some limits. Another limitation involves the lag time between harvest and regeneration. All of the current inventory results indicate increased harvesting that may have occurred in the latter years of the inventory periods. Because of this, it is possible that some of the harvested timberland was sampled between harvest and the development of regeneration. Regeneration assessment must be ongoing if the dynamism of southern forestry is to be described and major trends are to be identified.

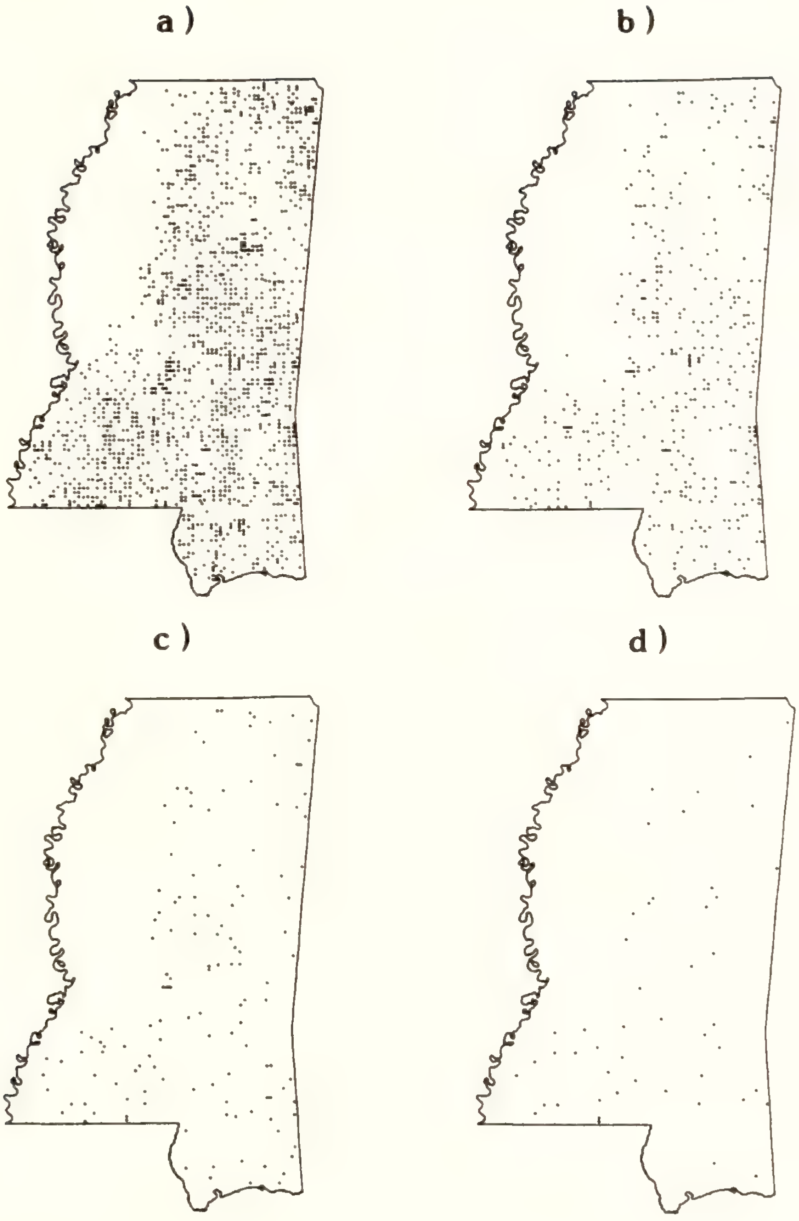


Figure 2. Spatial analysis of pine regeneration on nonindustrial private pine site timberland in Mississippi, 1978-87: a)All plots, b)Harvested plots, c)Harvested plots with medium or low pine stocking, and d)Harvested plots with medium or low pine stocking on high sites (capable of growing at least 120 cubic feet per acre annually).

REFERENCES

Burnett, H. 1987: The great drought of 1986. American Forests. May/June issue. pp. 22-25.

Cain, M.D. 1988: A low-cost technique for increasing pine volume production in mixed pine-hardwood stands. Research Paper SO-246. New Orleans,LA: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 9 p.

Cohen, M. 1988: Drought measure spells relief for stricken forest farmers. Forest Farmer, Vol. 48, No. 1, pp. 24-25.

Kaltenberg, M.C. 1984: Evaluation of regeneration sampling methods: a monte carlo analysis using simulated stands. In: New Forests for a Changing World: Proceedings of the 1983 Convention of the Society of American Foresters; 1983 October 16-20, Portland, OR: 27-34.

Lenhart, J. D., W.T. McGrath, and T.L. Hackett. 1988: Fusiform rust trends in East Texas: 1969-1987. Southern Journal of Applied Forestry, Vol. 2, No. 4, pp. 259-261.

McWilliams, W.H. 1989: Harvest activity and residual pine stocking on private lands of six South Central Coastal States. Southern Journal of Applied Forestry. (in press).

Phillips, D.R. and J.A. Abercrombie, Jr. 1987: Pine-hardwood mixtures - a new concept in regeneration on harvested lands. Southern Journal of Applied Forestry, Vol. 11, No. 4, pp. 192-197.

Stein, W.I. 1984a: Reforestation surveys: an overview. In: New Forests for a Changing World: Proceedings of the 1983 Convention of the Society of American Foresters; 1983 October 16-20, Portland, OR: 2-6.

Stein, W.I. 1984b: Fixed-plot methods for evaluating forest regeneration. In: New Forests for a Changing World: Proceedings of the 1983 Convention of the Society of American Foresters; 1983 October 16-20, Portland, OR: 20-26.

USDA, Forest Service. 1988: The South's fourth forest: alternatives for the future. FRR-24. National Technical Information Service, Springfield, VA. 512 p.

A Permanent Plot System Designed for Making Natural Resources Management Decisions in a Rapidly Urbanizing County

James M. Hill, Joshua L. Royte, Jil M. Swearingen,
Keith Van Ness, William Warncke, and John E. Hench.

Maryland-National Capitol Park and Planning Commission,
Department of Parks, Natural Resources Division, Natural
Resources Management Group, Chevy Chase, Maryland 20815

ABSTRACT

A permanent plot system based on a 1000 ft. (305 m) state plane grid network was initiated in 1988 within the Montgomery County, Maryland park system. Plots were located at grid intersections within large contiguous forested tracts in response to a need to implement gypsy moth (Lymantria dispar) Integrated Pest Management monitoring activities, and designed to inventory a variety of organisms in addition to the gypsy moth. Data generated from inventories will serve as the basis to evaluate, predict, and mitigate future impacts on natural resources under our stewardship. The purposes of the permanent plot system are to monitor gypsy moth population trends, stand condition, population parameters of floral and faunal components, and to predict impacts of development activities to plants and animals. Inventory data will also guide the selection of species composition and spacing pattern for a county-wide reforestation program.

INTRODUCTION

Permanent plots have been established in different forms and for different purposes since naturalists first became interested in documenting changes in floristics over time. Early plots were generally in a controlled environment, such as Mendel's genetics experiments. Other plots more extensively examine every aspect of forest biomass - every input, output and all their changes over time, as in the Hubbard Brook Experimental Forest (Siccama et al. 1970). The purpose of establishing plots may be solely for a repeatable, quantifiable inventory of what exists in a certain area at a certain time, e.g. an inventory documenting the vegetation of an old growth forest (Leopold et al. 1988). The inventory may be a tool for quantifying changes in forest growth over time in order to make prediction models such as JABOWA (Botkins et al. 1972, also see Elk et al. 1987). One of the longest running permanent plot systems in the country is conducted by the United States Forest Service. The goal of this national inventory system is to assess the nation's forest resource, to periodically report on changes in forest growth and changes in the use of forests (pers. comm. John Peters USFS 1989).

The Mongomery County, Maryland Department of Parks has initiated a permanent plot program for determining many of the above processes. As stewards of nearly 29,000 acres of park lands, it is imperative to know the condition and composition of this land in order to provide proper management. We have chosen the permanent plot method of inventory because it is quantifiable and repeatable, thus making it possible to reinventory plots for a measure of changes in time. As it is unlikely that any of the park's forests are over 80 years old, and still in a transitional or agradation phase with large amounts of tree mortality due to repeated gypsy moth defoliation, permanent plots will by useful in documenting compositional changes.

The Natural Resources Management Group (NRMG) established its permanent plot program for Montgomery County Parks in 1988. The objectives of this program are: 1) to systematically inventory flora and fauna, 2) to systematically monitor gypsy moth populations utilizing egg mass surveys, 3) to ascertain forest stand susceptibility to gypsy moth defoliation, 4) to document changes in forest floristics and physiognomy resulting from natural succession as well as gypsy moth defoliation, 5) to provide a standard method for the classification of vegetation associations, 6) to use vegetation classifications for resource evaluation, mitigation, acquisition, restoration and afforestation.

The MNCPPC's Permanent Plot Program is crucial to the successful implementation of the Department's gypsy moth suppression program, reforestation and mitigation needs, as well as the creation of a reliable base line inventory of the County's parkland. In short, the permanent plot program will help to ensure the preservation and maintenance of parkland in Montgomery County.

METHODS

Study Area

Montgomery County is bordered by Washington, D. C. to the south, the Potomac River to the west and the Patuxent River to the north. Physiographically, it is almost entirely within Maryland's Piedmont provence. Oak-hickory forests cover hilly upland areas, while beech-tulip poplar dominate the series of narrow stream valleys which drain the region. Ninety two percent of the county's soils are well drained to excessively well drained upland soils. Demographics range from urban in the south to suburban and rural in the northern half of the county. The population of Montgomery County is approximately 620,000 (Hench 1988).

Methods vary according to the ecology of the plants and animals being inventoried. The actual permanent plots may be ideal for vegetation and gypsy moth inventories but impractical for vertebrate inventories. The strand which connects all of our inventories is a 1000 ft. state planar grid. Each ecological or physical component that is studied can be referenced to a grid location and thus mapped. This becomes extremely important when master planning a park (see Natural Resources Management). The mapped layers of ecological or physical inventory data can be accurateley combined with exisiting mapped data to predict changes to the resource from proposed management actions. Eventually a Geographic Information System (GIS) will enable us to digitally relate plant, pest species and other animal surveys with physical parameters such as soils, slope, aspect, and hydrology.

Forest Pests Inventory

In the 1988-1989 season 714 gypsy moth egg mass survey points were established by the NRMG. Egg masses were counted in a 1/40th of an acre circle plot centered around each of the survey points. Survey points are located on the 1000 ft. grid intersections in the larger parks. In small neighborhood parks the survey points were located in the most susceptible area of the park. The survey points in stream valley parks are located every 1000 ft. along the stream valley, alternating between floodplain, mid-slope and upland areas. Initial tree surveys were conducted at each survey point using a 10 B.A.F. prism. Stand susceptibility was determined based on percent of favored host species present (e.g. oaks). This season (1989-1990) we will be able to concentrate our monitoring efforts on those areas with a high susceptibility and/or a current infestation of gypsy moths. By establishing permanent plots at the survey points we will simplify and standardize our monitoring program since plot boundaries are no longer estimated. Permanent plot vegetation inventories will help track defoliation and mortality of susceptible trees and stands and ascertain the effects of various levels of mortality on other forest vegetation and animal life.

A database is being set up for all of the Montgomery County parks that are impacted by gypsy moth outbreaks. This database will include plot numbers, acreage of forests, susceptibility of forest to outbreaks, annual determination of forest defoliation, mortality, numbers of egg masses per acre and type treatment the area received. Once established, this system will allow us to monitor and predict gypsy moth population trends and the efficacy of various treatment methods. Incorporation of this data into a GIS program will allow us to illuminate correlations between other baseline data and pest trends.

The current method of gypsy moth control in Montgomery County parklands consists of aerial spraying with the pesticide Bacillus thuringiensis (B.t.). The MNCPPC Integrated Pest Management program will consider other alternatives as well, including the application of selected host-specific viruses and fungi, the introduction of parasites and predators and other methods which are effective yet pose a minimum threat to non-target forest species. Having a permanent plot system in place will greatly facilitate in assessing the viability and success of various treatment methods used.

A variety of pest monitoring programs can be instituted at the permanent plots with minimum effort. The grid system and sampling points will already be established and baseline data on potential hosts for other pest species will be available.
Of current concern are various anthracnoses now causing considerable mortality to flowering dogwood (Cornus florida), and sycamore (Platanus occidentalis).

Vegetative Plots

Permanent plots are used as a sampling technique that is replicable by others in comparable forest types for the documentation of vegetative changes over time. We have chosen the plot size and methods in order to minimize time for plot inventory, (thus maximizing the sample size), and to maximize uniformity with permanent plot programs elsewhere in the country and abroad (EDAX 1989).

The plot size is 10 m x 10 m or 100 m^2 (approx. 1/40th ac.) with four sub plots, two for inventory of herbaceous layer vegetation, and two for measuring shrub layer vegetation. The plot boundaries are aligned with true north, south, east, and west with the grid intersection in the center. White plastic stakes, 40 cm long, are used to mark the SW, SE, and NE corners of each plot; a 1.75 m long steel post is used to mark the NW corner. Characterization of forest type, hydrology, and a visual estimate of canopy cover are recorded. Slope and aspect reflect an average for the entire plot. Data for soil type comes from Soil Conservation Service maps. Information on land use history is gathered from park historians, oral and photographic evidence.

In the 10 m x 10 m plot all living and dead stems over 2 cm dbh are counted, measured, assigned a plot specific reference number, and mapped. Fallen logs over 10 cm caliper are also mapped. The species, number, and size class [small (< 1 cm), medium (1-2 cm), large (> 2 cm)] of vines climbing on standing mapped stems are noted. Trees whose centers at base height fall on the plot boundary are included in the data. Canopy trees occurring just outside of the plot boundary, but judged to affect the floristics and/or physiognomy of the plot are also recorded.

Two shrub sub-plots (each 2.5 m x 2.5 m) are located in the NW and SE corners of the 10 m x 10 m plot. Within these sub-plots all woody vegetation less than 2 cm dbh is tallied in categories of greater or less than one meter high. Pokeweed (Phytolacca americana), which reaches shrub dimensions, is included in this tally. Two herb sub-plots (each 1 m x 1 m) are located in plot corners opposite the shrub sub-plots. In these, all herbaceous and woody vegetation less than 2 cm dbh and less than one meter high is tallied and a percent cover by species recorded. The percent cover and thickness of leaf litter and woody debris is recorded.

The initial plots sampled in this program have been in forests most susceptible to gypsy moth defoliation; dry upland forests dominated by oaks. A more extensive inventory will include floodplains, meadows, and specialty areas such as serpentine and shale barrens.

Data will be computerized using LOTUS IV, making it readily available for others as an ascii file. Vegetation community description, classification and ordination will be performed with statistical programs from SAS (Ray 1982) and DECORANA (Hill 1979).

Seed Bank

We have observed that plots located within forests suffering moderate to high mortality from gypsy moth defoliation exhibit a rapid release of plant species not normally found in mature shaded forests. One question that arose as a result of these observations was: Did these plants arise from an existing seed bank or were they transported there immediately following defoliation? A simple sampling scheme was developed whereby soil and leaf litter samples are sifted through two wire mesh sieves mounted on wood frames, their size being 1 m x 0.6 m, and constructed from 2.5 cm x 7.5 cm lumber. Samples are placed on the top sieve, the mesh size being 1/4 in. hardware cloth. The bottom sieve is window screen, and the two sieves are agitated together until the larger debris are sorted by the top sieve and discarded. Seeds are collected from the bottom sieve after further agitation.

Samples are collected from two 1 m^2 plots located on opposite sides on the outer edge of plot sides. The total volume of material sieved from each subplot is 8.6 liters.

Vertebrates

Methods for the following faunal groups are currently being considered: breeding passerine birds, hawks and owls, wintering birds, small mammals, furbearing mammals, and herptiles. Some of the methods being considered include 1000 foot by 150 foot belt transects for breeding passerine birds and active searching along grid referenced transects for herptiles.

Birds are proposed to be surveyed along transect lines between established permanent plots while walking at a slow pace. Stationary point counts are proposed to be conducted at each plot and at 100 m intervals along the transect, listening and watching for 3 minutes while turning 360 degrees (modified from Robbins 1978). Singing or drumming males, and calling or sighted individuals are mapped along each transect. All transects are run twice. Singing males mapped in the same location on both runs are considered probable breeding birds. Habitat breaks are noted along the transect.

The establishment of permanent scent-station transects is being tested to measure the abundance of fur bearing mammals. Scent-station indices can accurately reflect trends in the population of bobcats, raccoons, and gray foxes (Conner et al. 1983), and other fur bearing mammals as well (Humphrey and Zinn 1982, Sumner and Hill 1980). Methods are adapted from Roughton and Sweeney (1982). Station spacing as used by Conner et al. (1983), may not be applicable to our county park, and may be determined by use of home ranges of the species of interest.

Surveys for reptiles and amphibians may also be conducted along transects between plots. Suitable substrates such as logs, rocks and standing dead timber are searched along the route, and the number of log turns, rock turns, etc. are noted. Data are tabulated as e.g., the number of red-backed salamanders (Plethodon cinereus) per number of logs or rocks turned. All substrates are returned to their original position so that further surveys can be conducted as needed. Habitat or ecotype changes are noted along the transects.

Terrestrial Invertebrates

Surveys for invertebrates are centered on plot locations but are not necessarily limited to the confines of the plot. For example, during Spring 1989 samples of terrestrial insects were collected from foliage using beating sheets for a preliminary study of possible affects of an aerially applied insecticide used for gypsy moth suppression on non-target species. Samples were collected from a larger area centered on permanent plots. Plots served simply as a reference point in this case. In other groups of invertebrates, such as terrestrial gastropods, the plot might serve as the sampling unit, particularly if one goal of the sampling is to compare cover type or amount with the results of the faunal survey. Another method of invertebrate inventory we are considering involves fogging the forest canopy with insecticides such as pyrethrine (pers. comm. Terry Erwin 1989).

Rare Species

Since rare species are habitat specific and not as likely to occur within the confines of a plot as outside, the value of the plot work is more likely to be in accurately locating element occurrences, in reference to the grid intersections. However, since detailed work on plots may involve closer than usual examination of individuals, elements are sometimes identified from plots.

The degree of rarity of an element can be divided into four categories; Federal, State, County, and Local. Federal and State rare species are known from fewer populations and have fewer total individuals than species which are rare county-wide or in a particular local flora of the county, such as within a park. State or Federal rare species programs would not be concerned with species of County or local significance. However, natural resources and planning departments within counties should be concerned with tracking uncommon or local occurrences, particularly if the county is experiencing rapid or uncontrolled growth and a goal is to preserve the best remaining habitats or natural areas. The best way to judge whether a given area should or should not be developed is to gain an understanding of species richness and species quality. In most situations State or Federal species are not involved but those of county or local significance may be. Permanent plots near a rare species will be valuable in helping to describe the species habitat and in recognizing plant or animal associations.

Although not yet begun, the NRMG envisions utilization of permanent plots by outside volunteers with expertise in other specialized groups. Lichens, mosses, liverworts, fungi, slime molds, forest litter organisms, etc. could be sampled on a seasonal or annual basis within the plots. These specialists would benefit their studies by having access to data on physical and biotic parameters previously inventoried for each plot.

Natural Resources Management

The NRMG provides recommendation to park staff within all levels of the Department of Parks, the county Planning Department, and other agencies within county government. Management decisions and recommendations are based on a two-tiered approach to inventory of existing natural resources utilizing a coarse filter and a fine filter (Hench et al. 1986). During the coarse filter phase, a set of map sheets of uniform scale are generated detailing hydrology, drainage, wetland types (Cowardin et al. 1985), soils, topography, slopes, natural communities, and existing development. These are drafted on clear mylar and punch registered so that any or all of the plates can be overlaid and compared. In the future, these layers will be digitized into a GIS program to facilitate handling and manipulation. Areas that appear suitable for development based upon mapped physical features are then subjected to a fine filter.

During the fine filter, field work is performed at the proposed areas for possible development to determine what species would be impacted and to assure that no rare or unusual species would be severely impacted. Common trees herbs and shrubs are also inventoried (trees and shrubs are measured (dbh)), so they may be replaced either on or off site within the park system.

Prediction Modeling

Data gathered from permanent plots is required for daily management decisions regarding various natural resources programs in the park system, including gypsy moth suppression, mitigation for short and long-term impacts of development on or off site of parks, and reforestation or afforestation needs. By utilizing permanent plot inventory data gathered over time and subjecting it to statistical analysis, predictions for future conditions can be attempted and thereby brought into the management decision making process. Prediction models will be based on data from close monitoring of conditions at permanent plots at appropriate intervals over a number of years.

Prediction of Environmental Impacts

Prediction of the short and long-term impacts to sites from natural and human-caused disturbance can be attempted when comparisons are made to sites with similar physical and natural characteristics and histories. One goal in modelling using inventory plots is to lessen the need for extensive fieldwork at large sites and avoid the need for additional repeated sampling over one or more growing seasons. It is important to emphasize that the goal is not to replace field inspection.

Reforestation Requirements

A reforestation policy for the park system was recently proposed. The goal of this policy is to provide stocking and maintenance of trees and shrubs native to the piedmont of Maryland on unforested or underforested parkland. The purpose of reforestation or afforestation is for the creation of habitat for a high diversity of native plant and wildlife species and for the recreational enjoyment for park users. Planting is done on parkland where present features are no longer desireable, such as on existing agricultural fields, along streams that are mowed too near the banks, or other areas where excessive or unneeded mowing has occurred (e.g. steep slopes or poorly drained sites). Planning reforestation requires a prediction of the optimal combination of species and spacing patterns that would produce the healthiest, fastest growing, and most natural appearing forest for a given set of physical and biotic characteristics. The goal is not to maximize timber yield as in many forestry programs, but rather to increase the natural diversity and passive recreational opportunities of the site. Because of this approach conventional forestry guidelines for reforestation are not always applicable. Thus far we have relied on hand planting of 3-10 year old hardwood stock, mostly on recently mowed areas, utilizing variable spacing patterns averaging 3-4 m. The composition of plantings is reflective of forest types inventoried in permanent plots on similar sites.

Experimental Uses of Permanent Plots

In any area where manipulative experiments or management techniques alter the vegetation type, permanent plots will prove useful in providing a reliable reference of existing plant communities. Before and after manipulation the plots are reinventoried to examine how manipulation is affecting changes in tree, shrub, and herbaceous layer floristics. Plots established in areas experiencing forest mortality from gypsy moth defoliation will allow us to monitor how these areas respond to increased insolation, enrichment of the soil from leaf-fall, frass, and other factors. Studies of regrowth in these areas may be instructive in designing reforestation programs. We envision that permanent plots will be used to test the short and long term effects of prescribed burns applied to serpentine and shale barrens. Barrens in the county harbor rare plant species which are presently being overshadowed by mature trees. Elsewhere in the state, similar barren are much more open. We suspect that this condition was maintained in the past by periodic fire.

We are interested in the possible use of fire or herbicide to control local populations of exotic plants, e.g. Polygonum perfoliatum. Plots will allow us to determine the efficacy of control methods. Deer are becoming somewhat of a pest species in Montgomery County because of their extensive overgrazing. A set of sample plots will be established in fenced-in exclosure areas to determine what the long term effect of deer exclusion is on forest dynamics.

REFERENCES

Botkin, D. B., J. F. Janak, and J. R. Wallis. 1972. Rationale, limitations and assumptions of a northeastern forest growth simulator. IBM J. Res. Devel., **16**(2):101-116.

Conner, M. C., R. F. Labisky, D. R. Progulske. 1983. Scent Sation indices as measures of population abundance for bobcats, raccoons, gray foxes, and opossums. Wildl. Soc. Bull. **11**(2)146-152.

Cowardin, L. M., V. Carter, F. C. Golet, E. T. La Roe. 1985. Classification of Wetlands and Deepwater Habitats of the United States. FWS/OBS-79/31.

EDAX 1989. **A Catalogue of Ecological Data.** Yale School of Forestry and Environmental Studies. New Haven, CT.

Elk, A. R., S. R. Shifley, and T. E. Burk, eds. 1987. **Forest Growth Modelling and Prediction**. USDA For. Serv. Gen. Tech. Rep. NC-120.

Hench, J. E. 1988. Predicting some effects of landscape changes on wildlife. Unpublished Ph.D. Dissertation. University of Maryland. College Park. 113 pp.

Hench, J. E., K. Van Ness, R. Gibbs. 1986. Developemtn of a Natural Resource Planning and Management Process. from Integrating Man and Nature in the Metropolitan Environment. Proc. Natl. Symp. on Urban Wildl., Chevy Chase, Md., 4-7 November L. W. Adams and D. L. Leedy, eds. Published by Natl. Inst. for Urban Wildl., Columbia, MD.

Hill, M. O. 1979. DECORANA. A FORTRAN program for detrended correspondence analysis and reciprocal averaging. Section of Ecology and Systematics, Cornell University, Ithica, NY.

Humphrey, S. R. and T. L. Zinn. 1982. Seasonal habitat use by river otters and everglade mink in Florida. J. Wildl. Manage. 46:375-381

Leopold, D. J., C. Reschke, D. S. Smith. 1988. Old-Growth Forests of Adirondack Park, New York. Natural Areas Journal **8**(3):166-189.

Ray, A. A. 1982. SAS users guide: statistics. SAS Institute, Inc., Cary, NC.

Robbins, C. S. 1978. Census techniques for forest birds. pp. 142-163 in proceedings of workshop: Management of Southern Forests for Nongame Birds. USDA., For. Serv. Gen. Tech. Rep. SE-14.

Roughton, R. D. and M. W. Sweeney. 1982 Refinements in scent-station methodology for assessing trends in carnivore populations. J. Wildl. Manage. 46:217-229.

Siccama, T. G., F. H. Bormann, and G. E. Likens. 1970. The Hubbard Brook Ecosystem Study: Productivity, nutrients, and phytosociolgy of the herbaceous layer. Ecol. Monogr. **40**:389-402.

Sumner, P. W. and E. P. Hill. 1970. Scent-stations as indices of abundance in some furbearers of Alabama. Proc. Southeast Assoc. Fish and Wildl. Ag. 34:572-583.

INFOR_1: an Analysis Program for Permanent Forest Inventories with Systematic Sampling

Piermaria Corona
Head of the Silvicultural Section (1)

Agostino Ferrara
Head of the Forest Inventory and Land Survey Service (1)

(1) S.A.F. - Centro di Sperimentazione Agricola e Forestale
via di Casalotti 300, 00166 Rome, Italy

Authors are also, respectively, Adjunct Professor of Silviculture and
Adjunct Professor of Forest Inventory at Basilicata University - Italy

ABSTRACT

The complexity of a territorial environment, made up by innumerable interactive elements, makes it possible to guarantee an information level adequate to the need of territorial planning only by means of computer aided methodologies. The management of territorial information is furthermore simplified by adopting a structured design of computer management systems (modular structure), which enable us to correctly manage information and to preserve the updating flexibility of the whole system. A *module* of a Territorial Information System can be definited as a set of equipment, hardware, software and human resources which are able to manage information deriving from certain sources. Every *module* must be able to work independentely, but also to full interact with other modules by means of proper computer procedures and apposite hardware configurations. In this context, *INFOR_1* can be considered a program of the *module* dedicated to the management, processing and analysis of data collected by a permanent forest inventory with systematic sampling.

In this paper, a brief illustration of the aims of *INFOR_1* and a short review of its graphical and numerical output features are reported. An illustration of interactions among data collected by forest inventories and by other data sources within the territorial information system is also given.

INTRODUCTION

A forest inventory can be considered a territorial planning means: it does not only represent the general picture of the biologic and silvicultural characteristics of forest ecosystems, but also provides the basis for an articulate program of territorial data monitoring. In this sense, a forest inventory can perform its maximum effectiveness only when included in a project for permanent territorial data analysis, which integrates automatic cartography and updating methods, also by means of remotely sensed data, and which is provided with an efficient system for consultation and distribution of the information. In order to qualify the realization of forest inventories of some Italian Regional Administrations, the program *INFOR_1* , which can be considered a dedicated program allowing interactive management and analysis of data collected by permanent forest inventories with systematic sampling, was worked out (Ferrara et al., 1987, 1988a, 1988b).

Systematic sampling is suitable for assessing forest stands conditions. Sample units, in fact, are easily and objectively located on the ground and appear to be well representative of the realities investigated. Moreover, they allow the full integration of the other systems concerning territorial data collection and management (geographical data and remotely sensed data).

In this paper, a brief illustration of the aims of *INFOR_1* and a short review of its graphical and numerical output features are reported. An illustration of interactions among data collected by forest inventories and by other data sources within the territorial information system is also given.

THE INFOR_1 DEVELOPMENT

The Main Characteristics of a Regional Forest Inventory

The main factors which condition a regional forest inventory planning are:

- the definition of the purposes to be achieved and the problems to be pointed out;
- the definition of the balance point between the importance of statistical information and the need of territorial localization of the collected information;
- the definition of the possibility to get, in the future, further lathent information to the ones obtained from standard data processing;
- the necessity to update information in the mean or only in the long term;
- the definition of the technical potential of the Public Administrations to manage the forest inventory data.

Moreover, the priority requirement of the public contractors for conformity to the methodology adopted by the National Forest Inventory and for observation of some precise financial and temporal limits were devoted great attention in the above forest inventories carried out by S.A.F. (Società Agricola e Forestale, Rome - Italy), on behalf of some Italian Regional Administrations.

Therefore, the characteristics suggested for the realization of those regional forest inventories were:

- the setting up of a geographical information system to record the main topographical, hydrographical, road and administrative data;
- the mapping of forest stands information: this is to be integrated in the geographical information system in order to locate the information obtained from the inventory;
- the preparation of a systematic sampling scheme, based on permanent sample plots, geographically identified within territorial monitoring projects; according to the National Forest Inventory (ISAFA, 1983), a systematic sampling scheme with a square grid (1 km-long side), resulting from the thickening of the National survey grid (3 km-long side), was adopted;
- the application of a computer system for storing, processing, analyzing and distributing the information;
- the establishment of a permanent system to continuously update forest information (every 3-5 years).

The Structure of *INFOR_1*

The complexity of a territorial environment, made up by innumerable interactive elements, makes it possible to guarantee an information level adequate to the need of territorial planning only by means of computer aided methodologies.

The management of territorial information is furthermore simplified by adopting a structured design of computer management systems (modular structure), which enable us to correctly manage information and to preserve the updating flexibility of the whole system. A *module* of a Territorial Information System can be definited as a set of equipment, hardware, software and human resources which are able to manage information deriving from certain sources. Every *module* must be able to work independentely, but also to full interact with other modules by means of proper computer procedures and apposite hardware configurations. In this context, *INFOR_1* can be considered the program of the *module* dedicated to the management, processing and analysis of data collected by a forest inventory with systematic sampling. The modular structure of the S.A.F forest inventory system and territorial information management is presented in figure 1.

The program was worked out in a preliminary version, aiming mainly at developing methodologies and procedures of data processing and presentation, on a 200 series hp computer. At present, it is being worked out in its up-to-date version on VAX/VMS operating systems, to enhance the global processing capability and to get a full integration with the modules dedicated to the management of the other territorial information.

The structure of the main sections forming the program is presented in figure 2. It is essentially made up by:

- an interface to manage the dialogue between the computer and the user by means of interactive menus;
- the section concerning the input, updating, correction and storing of data;
- the section dedicated to the print of raw data and standard tables (obtained from arranged subroutines);
- the section for the management of interactive (non standard) output tables and graphics;
- the procedures for the connection and dialogue with the other modules of the system.

An example of the spreadsheet, on which input, updating and data correction are based, is reported in figure 3a .

Program Outputs

As previously underlined, the program can provide two different kinds of output tables: those adjusted for standard elaboration and those interactively defined by the user. The first kind of tables is made through special subroutines called by the program, whose structure and features can be differentiated according to the user's requirements. The above tables usually include different data recapitulation and synthesis types of greater interest and of more immediate and frequent consultation.

From an applicative point of view, more interesting are the tables (called *estimate tables*) created through the interactive interrogation of the data base, by which mean, total value and related statistical indicators of the main dendrometric variables (e.g. number of stems, stand basal area, stand volume, stand volume mean annual increment, total height curve, etc.) can be directly pointed out. Various administrative, topographic, pedologic and silvicultural characteristics related to each sample point permit the definition of the strata (as clusters of sample points):

- regional and provincial administration, mountain community, commune, land use restrictions, property
- elevation, aspect, slope, local relief, accessibility
- soil stoiness, soil depth, soil unit, rock outcrops
- dendrocenotic type, main and accessory species, stand evolution level, silvicultural system, stand mean age, main function (production, protection, reserved, etc.)

It is possible to change the number and the kind of these characteristics by simply adding, removing, trasforming or modifying a column in the data base (it is possible, for instance, add a column related to the tree crown depth or related to the key utilization species and so on, if it will be necessary in the future).

The interrogation of the data base allows, by the choice of one or more variables, the selection of clusters of sample plots. For each strata is also possible to obtain the distribution of the dendrometric values to be estimated according to the species and the diametric classes. It is possible, for instance, to

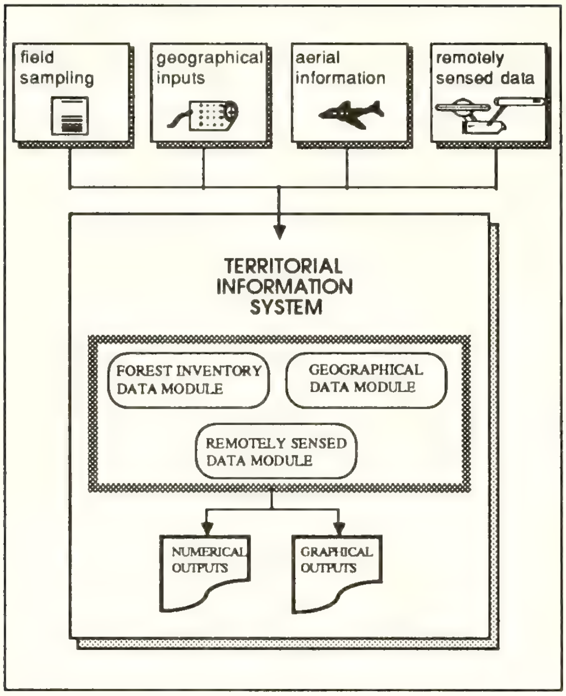


Figure 1. Components of an integrated information system, where each *module* self-administers a determined series of infomation, interacting, at the same time, with the remaining modules.

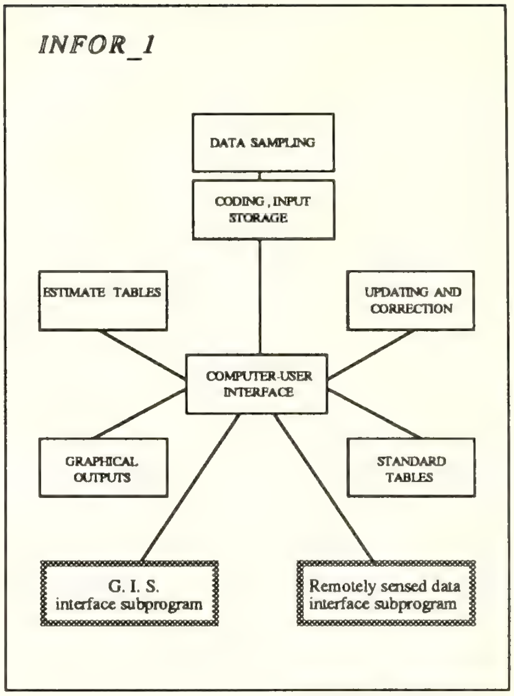


Figure 2. Structure of the program main sections.

estimate the *basal area* (or any other dendrometric variable listed above) of *chestnut coppices* (silvicultural characteristic) between *600 and 800 m a.s.l.* (topographic characteristic) in the *Province of Rome* (administrative characteristict). Therefore, it is possible to deeply examine many aspects concerning the inventoried stands by only changing the kind, the order and the number of the selected characteristics.

The estimate tables are set up according to an unique output structure, so that their reading became quickly friendly. For instance, the estimate of the stand volume in the administrative district of Bergamo, with respect to stand accessibility and elevation, is wanted: an example of the table outputted in relation to such an interrogation is presented in table 1. The table identification code [1] is reported on the upper left side, while the dendrometric variable [2], to which the estimate is referred (in this example, stand volume), is reported on the rigth side. In the following line, on the top of the respective columns, the coded names of the variables used for the interrogation of the data base are firstly presented:
- Province [3], i.e. administrative district;
- accessib. [4], i.e. accessibility degree (in this case divided into three classes, * = not accessible, ** = scarcely accessible, *** = accessible);
- elevation [5], i.e. elevation m a.s.l. (in this case divided into five classes).

On the same line are then presented:
- surface [6], i.e. forest stand total surface related to each of the selected strata;
- mean value [7] and related standard error [8];
- total [9], i.e. the total value of the selected variable in the selected strata.

Another example is illustrated in table 2.

Table 1. Example of table obtained by interactive interrogation of the data base.

ESTIMATE /29/12/25 [1]

PROVINCE OF BERGAMO : Total Stem Volume, m^3 [2]

selected characteristics						
Province [3]	accessibil. [4]	elevation [5]	surface (ha) [6]	mean/ha [7]	se [8]	total [9]
Bergamo			61400	146.031	4.332	8966464
	*		23400	140.358	6.570	3284376
		<=400	2900	122.131	16.750	354179
		401-800	10600	120.636	9.082	1278746
		801-1200	7800	158.401	11.406	1235525
		1201-1600	1900	203.864	27.587	387342
		>1600	200	142.917	80.417	28583
	**		10000	166.427	13.086	1664267
		<=400	1000	77.542	16.427	77542
		401-800	2900	163.358	16.389	473738
		801-1200	3000	187.011	25.994	561033
		1201-1600	2600	196.037	32.743	509696
		>1600	500	84.517	31.449	42258
	***		28000	143.494	6.160	4017822
		<=400	200	91.000	45.667	18200
		401-800	7200	130.349	10.778	938509
		801-1200	10800	139.377	9.452	1505267
		1201-1600	8800	164.364	12.363	1446404
		>1600	10000	109.442	32.389	109442

INFOR_1 - Bergamo

[1-9] see text

Table 2. Example of table obtained by interactive interrogation of the data base (*).

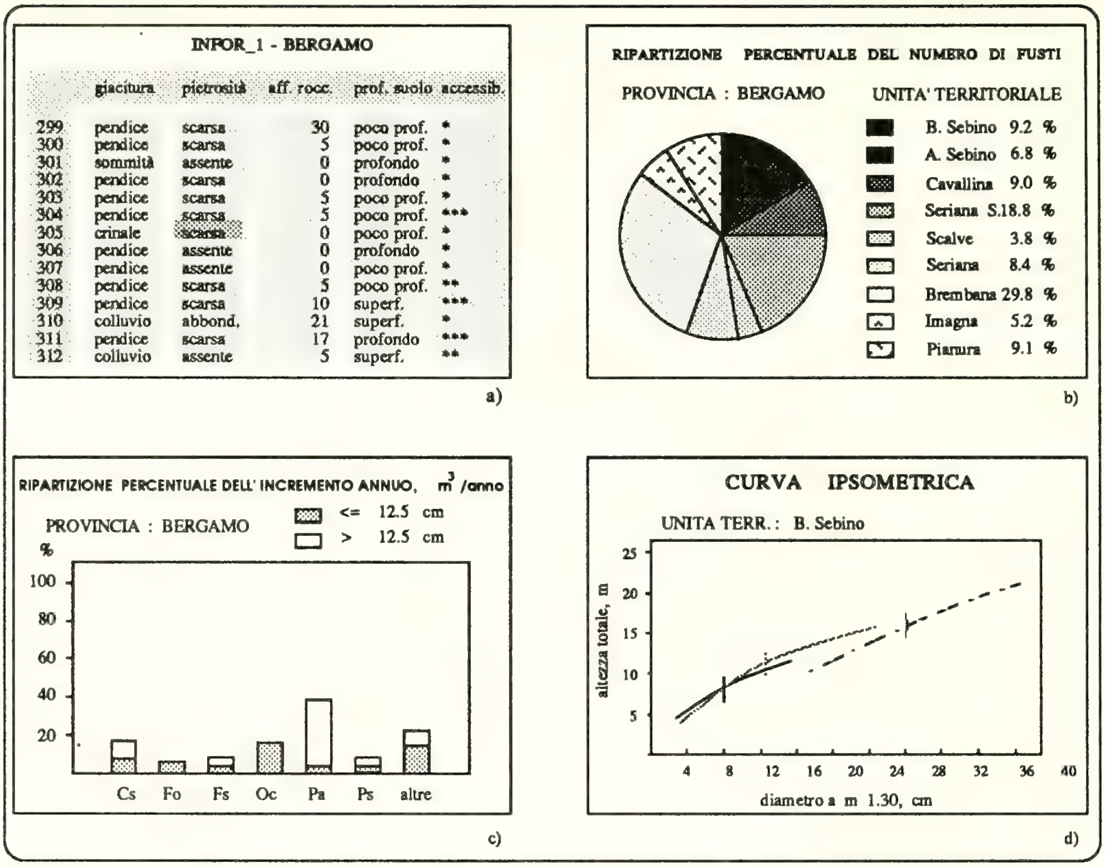
ESTIMATE /29/31

PROVINCE OF BERGAMO : Total Stem Volume, m^3

selected characteristics					
Province	silv. system	surface (ha)	mean/ha	se	total
Bergamo		61400	146.031	4.332	8966464
	coppices	34700	113.583	4.421	3941347
	in conversion	5300	156.970	11.043	827313
	high forests	21400	196.159	8.729	4197805

INFOR_1 - Bergamo

(*) The table may have a different length according to the choice made.



INFOR_1 - BERGAMO

	giacitura	pietrosità	aff. rocc.	prof. suolo	accessib.
299	pendice	scarsa	30	poco prof.	*
300	pendice	scarsa	5	poco prof.	*
301	sommità	assente	0	profondo	*
302	pendice	scarsa	0	profondo	*
303	pendice	scarsa	5	poco prof.	*
304	pendice	scarsa	5	poco prof.	***
305	crinale	scarsa	0	poco prof.	*
306	pendice	assente	0	profondo	*
307	pendice	assente	0	poco prof.	*
308	pendice	scarsa	5	poco prof.	**
309	pendice	scarsa	10	superf.	***
310	colluvio	abbond.	21	superf.	*
311	pendice	scarsa	17	profondo	***
312	colluvio	assente	5	superf.	**

Figure 3. Examples of the input, updating and data correction spreadsheet a) and of some graphical output features b), c), d).

All the values obtained by this way can be also presented by several types of graphical outputs (pies, istograms, bar charts etc.): some of the possible graphic outputs are illustrated in figure 3 b),c),d).

The integration of *INFOR_1* with the other system *modules.*

The merging of the information collected by *INFOR_1* within a wider context of territorial data analysis is accomplished by the integration of the *INFOR_1 module* with those dedicated to the management of geographic and remotely sensed data.

The *module* for the geographical data manipulation is based on a Tektronix graphic station operating on VAX 8350 and equipped with the proper devices (digitizer, plotters, printers, etc.). The *module* for remotely sensed data manipulation is made up by two Remote Sensing Machines with additional devices (CCD video camera, electrostatic plotter, freeze-frame, etc.) for the processing of images from LANDSAT and SPOT satellites. The use of different equipments in the three *modules* is fundamentally complementary. Actually, the geographical *module* and forest inventory *module* are an essential basis of information for deep forest data insight; on the other hand, remotely sensed data can be also profitably applied for updating forest inventories, forest maps, etc., because of its low cost, high informative content and possibilities of repetitiveness.

Example of integrations among different sources of territorial data are illustrated in figure 4 and 5.

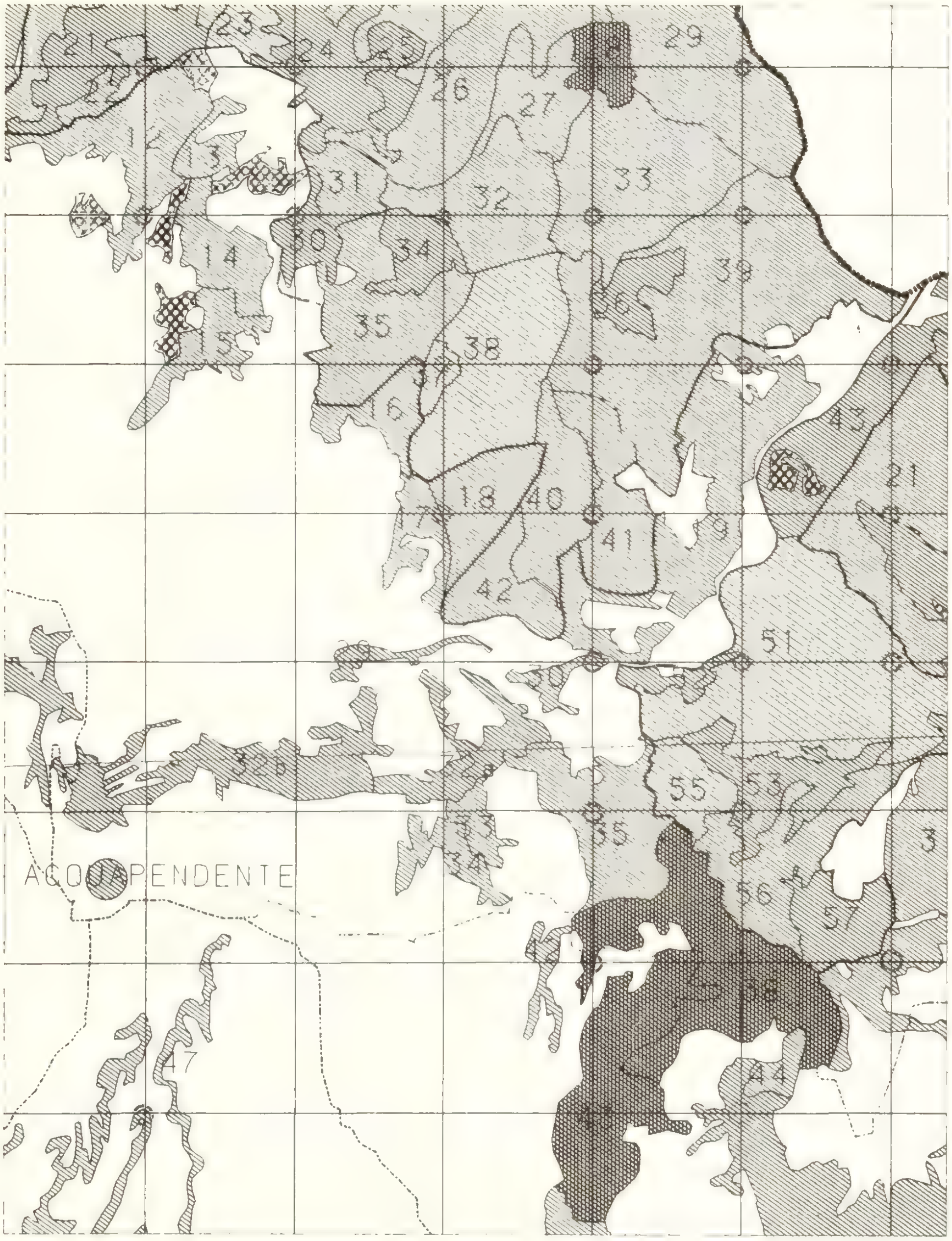


Figure 4. Example of integration between the data obtained by point sampling and forest mapping. The data bases of information collected through the two inquiry sources (data collected either from the points or from the forest stand compartments) are fully compatible and allow interactive interrogation and cross data elaboration.

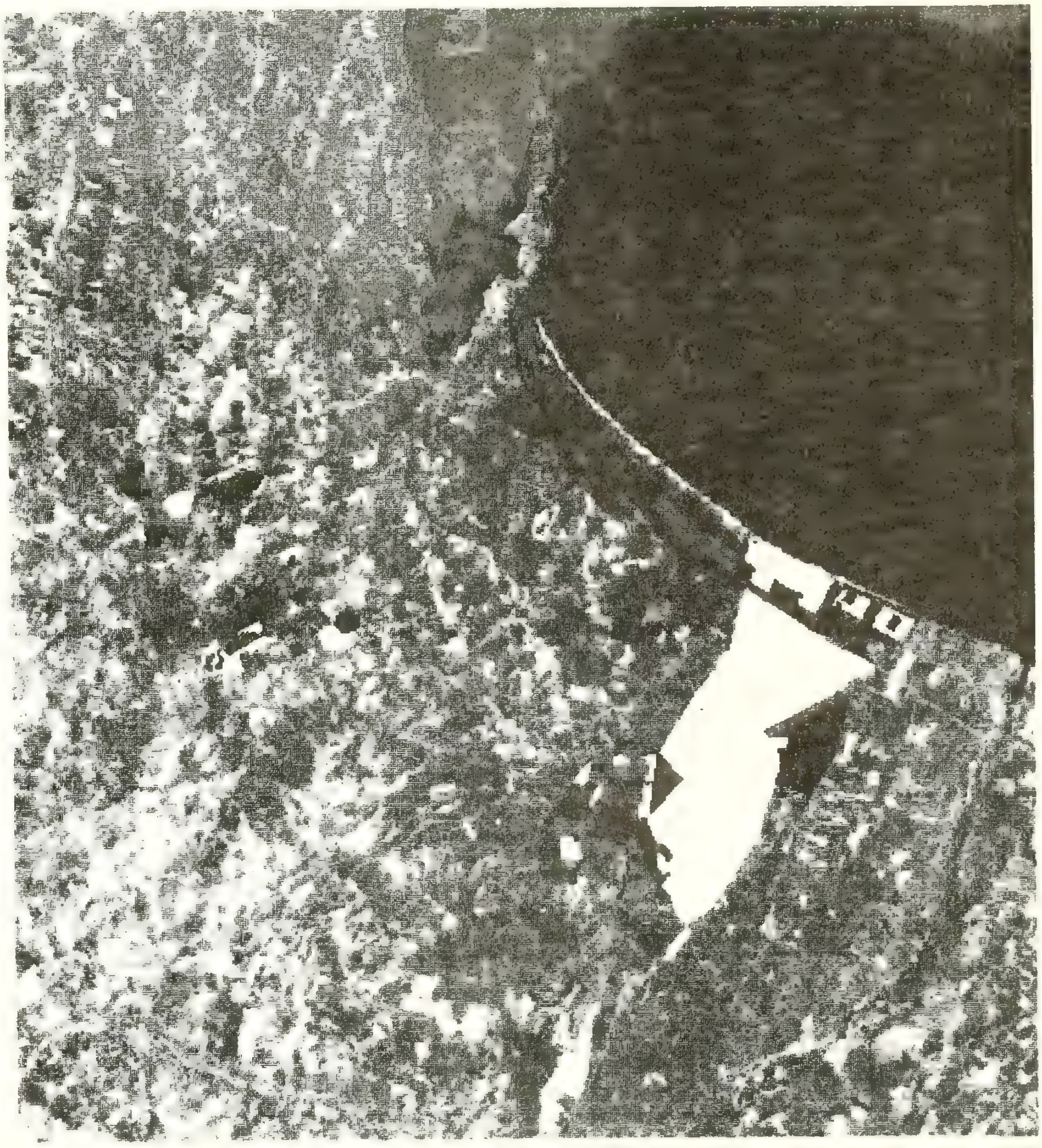


Figure 5. Example of full integration between raster and vectorial information.

CONCLUSIONS

The computer technology offers the means for trasforming the territorial data conventional analysis into an integrated system of territorial data management, properly structured by manipulation both of graphic and alphanumeric information. The integration of field forest data, remotely sensed forest information and cartographic supports allows a synergic outcome, whose value is higher than the simple sum of the partial outcomes, in relation to the potential of cross-checking and comparison, so that permanent forest inventories management systems attain their maximum information potentiality just within such a territorial information modular structure.

REFERENCES

Ferrara A., Corona P. et al., 1987: Indagine conoscitiva sui boschi della provincia di Varese: inventario e carta forestale. Regione Lombardia, Azienda Regionale delle Foreste, Milano

Ferrara A. et al., 1988a: Indagine conoscitiva sui boschi della provincia di Bergamo: inventario e carta forestale. Regione Lombardia, Azienda Regionale delle Foreste, Milano

Ferrara A. et al., 1988b: Indagine conoscitiva sui boschi della Parco Naturale Regionale dei Monti Simbruini (prima fase del piano di assestamento forestale): inventario e carta forestale. Regione Lazio, Assessorato Agricoltura e Foreste, Roma

ISAFA, 1983: Inventario Forestale Nazionale Italiano, progetto operativo. Trento

Inventory to Support Silvicultural Decisions

Susan L. Stout
Research Forester
USDA Forest Service, Northeastern Forest Experiment Station
Forestry Sciences Laboratory, P.O. Box 928, Warren, PA 16365

ABSTRACT

The SILVAH (Silviculture of Allegheny Hardwoods) system of stand inventory, analysis, and prescription facilitates the use of silvicultural research results. Timber management research in the Allegheny region of the Eastern United States has identified 14 stand characteristics that lead to a unique silvicultural prescription for an individual forest stand. An inventory scheme to determine these characteristics is one component of the SILVAH system. A set of computer programs has been developed to summarize inventory data and identify the prescription. Such a system reduces the subjectivity traditionally used in making silvicultural decisions, increases the consistency of silvicultural decisions from stand to stand and among silviculturists, and ensures that all key stand characteristics are assessed and considered in decision making. The automated identification of the "textbook" prescription for any stand provides support to the silvicultural decision-making process, and allows the silviculturist to focus on unique characteristics or special management considerations.

INTRODUCTION

The SILVAH system is built around a series of decision flow charts (Figures 1a through 1e) that indicate appropriate treatments for particular forest stand conditions. The system is designed for practicing foresters whose primary goal is high-quality sawtimber production, but includes options and modifications for coordination of wildlife, esthetics, recreation, and other forest uses where appropriate. To use the decision flow charts, it is necessary to make a stand inventory, and then to analyze the data collected to determine present stand conditions and the potential for future growth or regeneration. The SILVAH system in its entirety and many of the research results on which it is based are described elsewhere (Marquis and others 1984, Marquis and others in preparation). A computer implementation of the SILVAH system is also available (Marquis and Ernst in preparation).[1] This paper will focus on the relationship between inventory data and the silvicultural decision making process in the SILVAH system.

The particular guidelines shown in the decision charts are based primarily on research from a 6- or 7-county area in northwestern Pennsylvania, supplemented with results from northern New Hampshire, northern New York, north central West Virginia, and the central states. The guidelines are intended for use in the cherry-maple, beech-birch-maple, and oak-hickory forests of the Allegheny Plateau and Allegheny Mountain sections of Pennsylvania, New York, Maryland, and West Virginia. Testing and application has been extensive in the Allegheny region over an 8- or 9-year period. Caution is needed when applying these specific numerical guidelines outside the Allegheny region. The basic procedures and framework of the decision making process, however, are universally applicable. If local silvical knowledge is sufficient to permit the needed modification in decision variables, the system of stand inventory, analysis, and prescription can be adapted to any region.

---

[1] For additional information, write to the Forestry Sciences Laboratory at the address above.

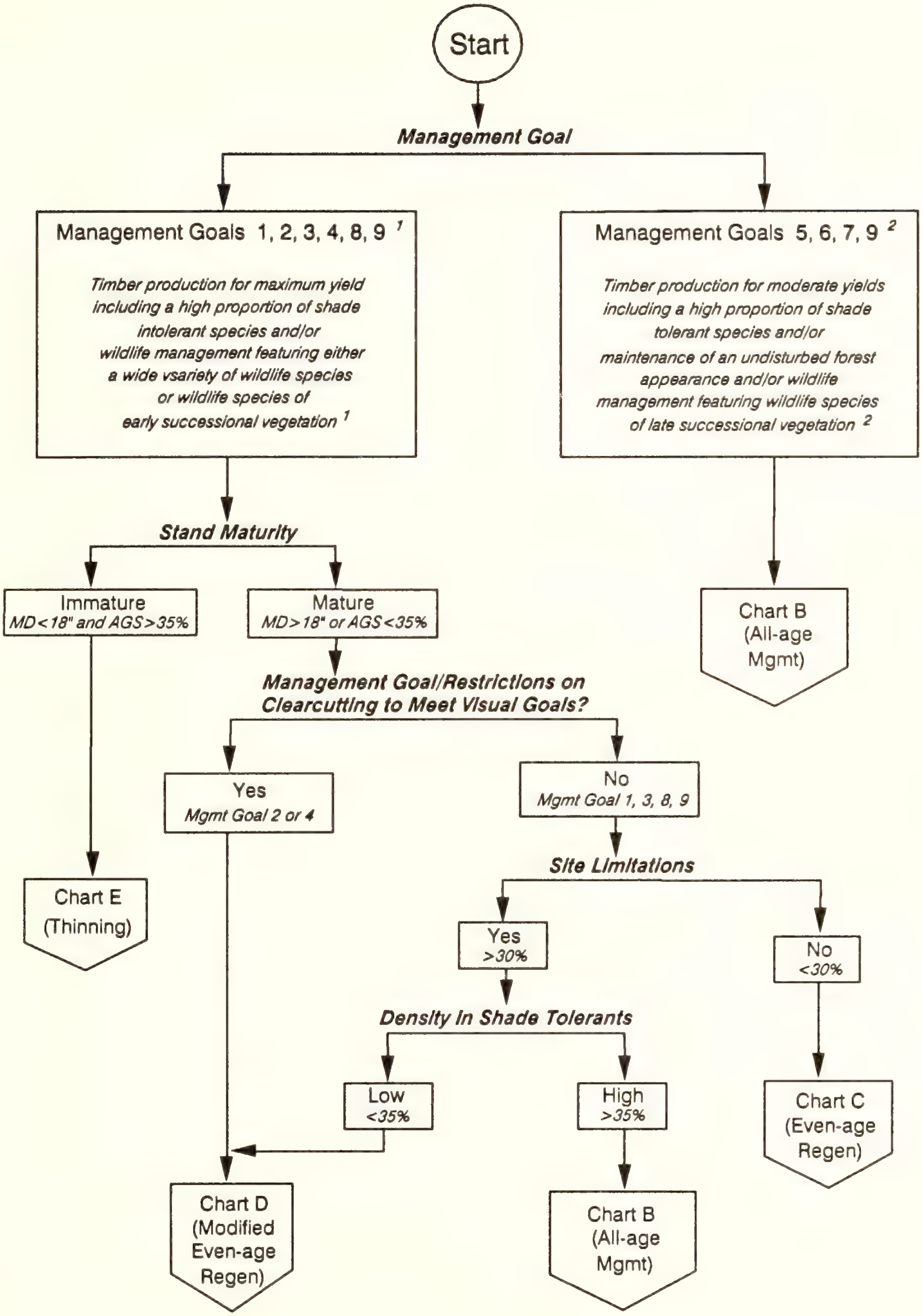

Start
Management Goal
Management Goals 1, 2, 3, 4, 8, 9 1
Timber production for maximum yield including a high proportion of shade intolerant species and/or wildlife management featuring either a wide vsariety of wildlife species or wildlife species of early successional vegetation 1
Management Goals 5, 6, 7, 9 2
Timber production for moderate yields including a high proportion of shade tolerant species and/or maintenance of an undisturbed forest appearance and/or wildlife management featuring wildlife species of late successional vegetation 2
Stand Maturity
Immature
MD<18" and AGS>35%
Mature
MD>18" or AGS<35%
Chart B
(All-age Mgmt)
Management Goal/Restrictions on Clearcutting to Meet Visual Goals?
Yes
Mgmt Goal 2 or 4
No
Mgmt Goal 1, 3, 8, 9
Chart E
(Thinning)
Site Limitations
Yes
>30%
No
<30%
Density in Shade Tolerants
Low
<35%
High
>35%
Chart C
(Even-age Regen)
Chart D
(Modified Even-age Regen)
Chart B
(All-age Mgmt)

Figure 1. Decision charts for Allegheny hardwood silvicultural decision support. (a) Master chart

Chart B. All-age Management

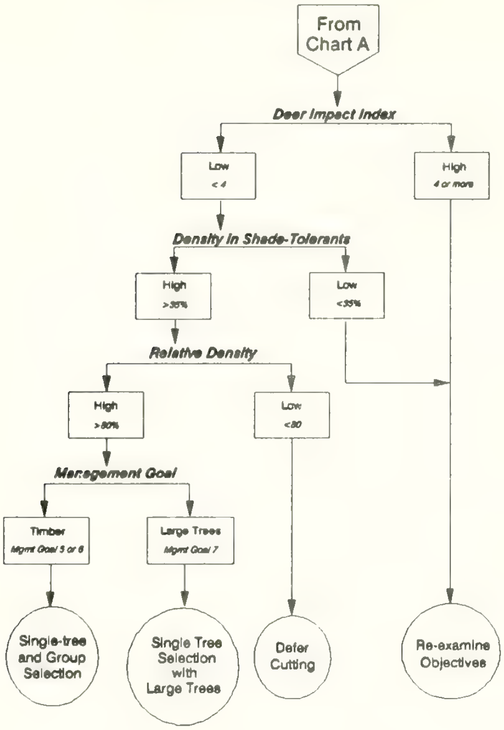


Chart C. Even-age Regeneration

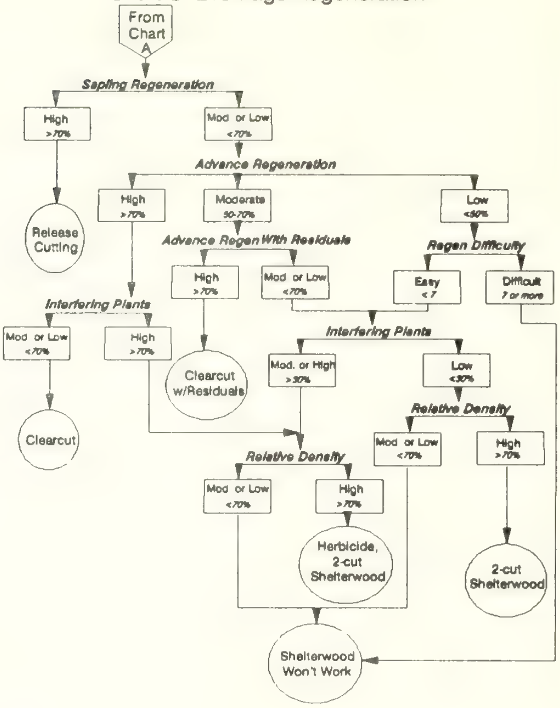


Chart D. Modified Even-age Regeneration

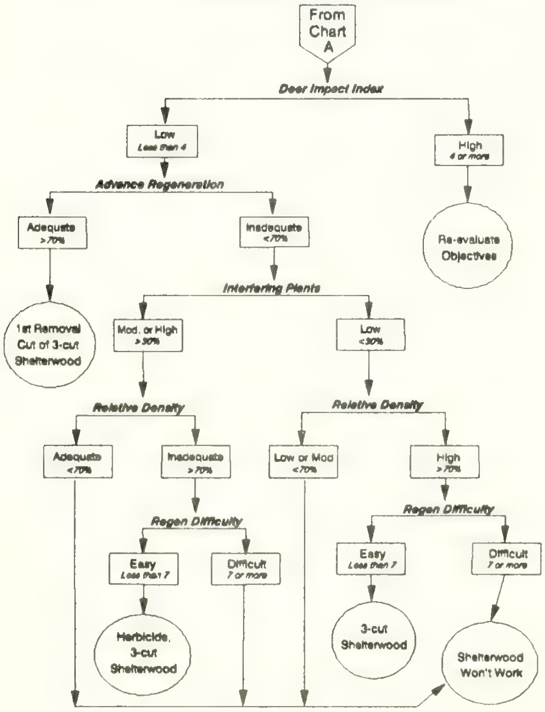
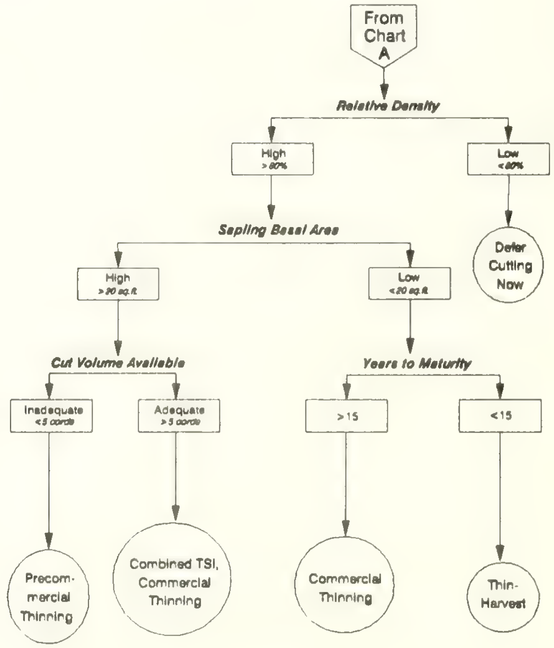


Chart E. Thinning



Figures 1b–e. Silvicultural decision charts for hardwood stands of the Alleghenies.

An important characteristic of the SILVAH system is that research results have been summarized into numerical guidelines that take advantage of traditional inventory data. For example, research has shown that many stands of the Allegheny region (Marquis 1975) are multi-aged, resulting from varying sequences of incomplete and nearly complete harvest cuts. Despite their irregular age class distribution, however, most stands respond productively to even-age management techniques. Average tree diameter is a good indicator of maturity. So, a definition of maturity has been developed that is based on an average diameter, easily calculated from a stand table. The stand table, in turn, can be developed by any conventional inventory procedure. When a conventional overstory inventory is accompanied by a simple, checkmark-based inventory of understory conditions, practicing foresters can easily calculate all the values needed to implement research results.

DECISION VARIABLES

Fourteen variables are assessed in the SILVAH decision-making process. These values are shown in bold italics in Figure 1. They can be divided into three subgroups, based on the source of the value. Site and Management Variables are determined by managers, rather than by inventory results. Overstory Variables are characteristics of the overstory stand, that is, all trees 1.0" in diameter breast height (dbh) and larger. The threshold diameter of 1.0" is important in the calculation of relative density and sapling basal area. Understory Variables are derived from a separate inventory procedure based on 6-foot radius plots.

Before conducting an inventory, it is not possible to determine which values will be needed, so the inventory procedure must supply data to allow calculation of 12 of the 14 values.

Management and Site Variables

These two values are not calculated from inventory data. They represent the primary way in which aesthetic, recreation, and wildlife habitat constraints and objectives are integrated with timber considerations in arriving at a prescription.

Management goal (Fig. 1a,b). Users of the SILVAH system specify a coded management goal for each individual forest stand, incorporating timber management, wildlife habitat, and aesthetic or recreational use considerations. This management goal is set by the landowner from among the alternatives provided. The range of alternatives is shown in Figure 1a. By specifying the appropriate management goal, even-aged harvest methods can be excluded from the range of prescriptions considered for a particular stand, or modifications to even-age harvests can be imposed. The reasons for these limits may be the tree or animal species sought, or aesthetic preference.

Deer impact index (Fig. 1b,d). At present, a knowledgeable forester must integrate what is known about available deer food in adjacent farm and forest land with any indications of the size of the local deer herd in order to project the pressure that deer will exert on regeneration development in the stand. Local herd size may be estimated by pellet group counts or by assessing browsing damage in this or adjacent harvested stands, or estimates provided by the appropriate game management agency may be used. Ultimately, we hope to calculate this value from stand inventory data.

Overstory Variables

Seven overstory factors are required to identify a prescription. All of these can be calculated from a complete stand table by species and diameter at breast height (dbh), and a rough estimate of tree quality. The quality estimate is an integrated assessment of risk and potential merchantability, with two classes. Acceptable growing stock (AGS) trees are those that will survive for at least 15 years and will contain a merchantable product in the future, if provided with good growing conditions. Unacceptable growing stock (UGS) trees are those that will not survive or trees of such poor form that they must be removed in any cutting treatments.

Many of the overstory variables are expressed as percentages of relative density. Because the forest types of the Allegheny region are stratified mixtures of many species, absolute measures of density, such as stems per acre or basal area per acre, are misleading indicators of some important stand conditions. Average maximum density (Ernst and Knapp 1985) in this type, for example, can vary by 50% in stands of the same average tree size, simply as a function of species composition. Relative density is determined with a tree-area ratio equation applied to the stand table by species group and diameter (Stout and Nyland 1986).

In general, a 10-factor prism inventory of all trees with a dbh of at least 1.0" is used, but any inventory procedure that will yield a stand table is acceptable. The SILVAH computer programs will print standard error estimates for basal area per acre if fixed- or variable-radius plot sampling is used to estimate the stand table, and warn users to regard prescriptions based on overstory sampling with a standard error in excess of 10% with caution. The program will also print a warning (regardless of the sampling procedure used) when a key variable is estimated to be close to a decision threshold.

Stand Maturity (Fig. 1a). Studies of individual tree financial maturity (Grisez and Mendel 1972, Mendel and others 1973) plus unpublished results of computer simulator runs in Allegheny stands of varying species composition have led to a rule of thumb that stands can be considered for final harvest sequences when the diameter at the mid-point of the basal area distribution of merchantable trees reaches 18". Stands with less than 35% relative density in AGS are also considered mature, regardless of the average diameter of the stems. It is more profitable to start over in such stands than to devote the site to production of low-value material. Both average diameter and relative density in AGS can be calculated from the stand table.

Density in Shade Tolerants (Fig. 1a,b). When management goals dictate the selection of all-age silvicultural prescriptions for a stand, the range of stand conditions for regeneration is narrowed to those favorable for shade-tolerant species. The shade-tolerant species of the Allegheny region forest are not prolific seeders, and a minimum component of these species must be present to ensure successful regeneration required under all-age systems. When this component is present, all-age silviculture may also be prescribed for sites on which poor drainage or excessive surface stoniness would otherwise limit the regeneration capability of the stand. This value is a subtotal of the relative density calculation.

Regeneration Difficulty (Fig. 1c,d). This value is the sum of the Deer Impact Index and a Seed Source Index calculated from the overstory stand table. Information about seed production by important Allegheny and northern hardwood species has been codified to allow calculation of the adequacy of seed likely to be produced (Marquis and others 1975), and this value is combined with the deer pressure to yield an integrated assessment of the difficulty of regenerating the stand.

Relative Density (Fig.1c,d,e). The total relative density of the overstory provides an estimate of the light conditions on the forest floor. If overstory density is below a threshold level and seedlings are not developing, there is likely to be some other problem. Similarly, overstory relative density values define the boundaries of the management zone (Ernst and Knapp 1985) to be maintained by intermediate cuttings. The boundary levels are a compromise among stand and individual tree growth, tree quality, and understory development, and represents the lower level of the management zone. Relative density is calculated for each species group and size class, then summed across the stand table.

Sapling Basal Area (Fig. 1e). Very shade-tolerant species will often persist in the lower strata of forest stands in the Allegheny region for as long as 50 or more years of stand development, and remain capable of responding to release. If they are ignored in intermediate treatments, they may command larger and larger proportions of the growing space without ever reaching merchantability. If the basal area of this strata exceeds a threshold level, intermediate treatments must either be postponed or must include some removals in this strata, usually at a net cost to the landowner. The value is a subtotal of the basal area stand table.

Cut Volume Available (Fig. 1e). In most areas, commercial operators require some minimum yield from intermediate treatments before they will consider bidding on the associated sale. This value is estimated from the stand table of species and diameter using regional equations for merchantable height. Optionally, the inventory may include estimates of merchantable height, factory log-grade of the first log, and cull percentage. If a silviculturally desirable commercial thinning will not yield this threshold volume, a precommercial thinning is prescribed.

Years to Maturity (Fig. 1e). For stands under even-age silvicultural systems, the cutting strategy for intermediate treatments near the end of the rotation is modified to capitalize on the value of some of the individually mature trees. Since maturity is defined as the time when the diameter at the mid-point of the basal area distribution of merchantable size trees reaches 18", estimates of the years left in the rotation can be based on the projected increase in this mid-point diameter. The calculation is based on an average growth rate weighted by species composition. Thus, years to maturity can be calculated from the stand table.

Understory Variables

Five additional variables are assessed from an inventory of understory conditions. In the forest types covered by the SILVAH system, advance regeneration is the key to regeneration success (Marquis and others 1975, Marquis and Bjorkbom 1982). The amount and distribution of advance regeneration are important, and the amount and distribution of interfering plants and site limiting conditions also are used in determining a silvicultural prescription. Thus the sampling procedure used is the assessment of a series of 6-foot radius plots and their surrounding area against a series of criteria. If a plot meets a criterion, such as some minimum number of advance seedlings or density of interfering plants, that criterion is checked off on the inventory tally sheet. The decision flow charts are based on the proportion of plots that met the criteria. Detailed descriptions of the criteria for each understory inventory category are found in Marquis and others (1984).

Site Limitations (Fig. 1a). Poor soil drainage or excessive surface rockiness can threaten the survival of advance seedlings in the harsh conditions following complete overstory removal under even-age systems. The understory inventory summary will indicate whether this condition prevails over more than 30% of the stand area, in which case restrictions on even-age harvests must be imposed.

Sapling regeneration (Fig. 1c). If a large proportion of the sampled understory plots is stocked with sapling-size trees of a distinctly different age class than the overstory, this dense layer will preclude development of new advance regeneration. It is the de facto next stand. Note that in the conventional SILVAH inventory, this value is determined from the fixed-area small plots of the understory cruise, rather than the prism cruise of the overstory. This ensures both better sampling of the small trees and an estimate of the uniformity of their distribution.

Advance Regeneration (Fig. 1c,d). This is the proportion of plots stocked with seedling-size advance regeneration of acceptable species. In oak stands, the proportion is supplemented with a prediction of the stocking to be provided by oak stump sprouts (Sander and others 1976), which is calculated from the overstory stand table by species and diameter class.

Advance Regeneration Plus Residuals (Fig. 1c). Sapling to pole-size stems of shade-tolerant species and very high bole and crown quality can be left at the time of a harvest cut to form a part of the new stand (Marquis 1981). If the level of advance seedling regeneration is moderate, regeneration plots stocked with such trees can be added to those stocked with advance seedlings in determining whether a harvest is appropriate.

Interfering Plants (Fig. 1c,d). Many woody and non-woody plants interfere with the development of desirable regeneration under mixed hardwood stands (Horsley 1977a, 1977b, 1983, 1986). The percentage cover and distribution of these plants are estimated from the inventory. Threshold levels trigger prescriptions including herbicide treatment; these levels vary with the amount of advance regeneration detected in the inventory.

DISCUSSION

Often, research results are reported in terms that are difficult for practicing foresters to integrate into daily operations. By codifying research results from the Allegheny region into a form easily assessed through conventional inventory procedures, and packaging them in a computer program, Allegheny region researchers have increased the use of research results. Guidelines on stand maturity, harvest methods including removal of interfering understories and modifications to accommodate site limiting conditions, intermediate treatments, and all-age silviculture are incorporated into a series of decision flow charts that can be traced to a unique prescription for any stand, based on data from a two-part inventory and coded management constraints. The first part of the inventory is any conventional overstory inventory that uses a threshold dbh of 1.0" and includes data on species, diameter, and a rough assessment of tree quality. The second part of the inventory is a checkmark tally of 6-foot radius plots on which a variety of key understory plant and site conditions are assessed. Results of these two inventories are summarized to give 14 values for any stand, and these values, combined with the management goal for the stand and an estimate of deer pressure, will indicate the prescription supported by silvicultural research. The system is used as an aid to professional silvicultural decision-making throughout the Allegheny region.

REFERENCES

Ernst, R. and Knapp, W.H., 1985:Forest stand density and stocking: concepts, terms and the use of stocking guides. USDA Forest Service, General Technical Report WO-44, 8 p.

Grisez, T.J. and Mendel, J.J., 1972:The rate of value increase for black cherry, red maple, and white ash. USDA Forest Service, Northeastern Forest Experiment Station, Research Paper NE-231, 26 p.

Horsley, S.B., 1977a:Allelopathic inhibition of black cherry by fern, grass, goldenrod, and aster. Canadian Journal of Forest Research, vol. 7, no. 2, pp. 205-216.

Horsley, S.B., 1977b:Allelopathic inhibition of black cherry II. Inhibition by woodland grass, ferns, and club moss. Canadian Journal of Forest Research, vol. 7, no. 3, pp. 515-519.

Horsley, S.B. 1983:Interference with desirable northern hardwood regeneration by herbaceous and woody plants. In Finley, J., Cochran, R.S., and Grace, J.R., eds. Regenerating hardwood stands: Proceedings. 15-16 March 1983, University Park, PA. The Pennsylvania State University School of Forest Resources, Cooperative Extension Service, pp. 81-93.

Horsley, S.B., 1986:Evaluation of hay-scented fern interference with black cherry. American Journal of Botany, vol. 73, pp. 668-669.

Marquis, D.A., 1975:The Allegheny hardwood forests of Pennsylvania. USDA Forest Service, Northeastern Forest Experiment Station, General Technical Report NE-15, 32 p.

Marquis, D.A., 1981:Removal or retention of unmerchantable saplings in Allegheny hardwoods: effect on regeneration after clearcutting. Journal of Forestry, vol. 75, no. 5, pp. 280-283.

Marquis, D.A., ed., in preparation:Quantitative Silviculture for hardwood stands of the Alleghenies.

Marquis, D.A. and Bjorkbom, J.C., 1982:Guidelines for evaluating regeneration before and after clearcutting Allegheny hardwoods. USDA Forest Service Northeastern Forest Experiment Station Research Note NE-307, 4 p.

Marquis, D.A., and Ernst, R.L., in preparation:User's Guide to the SILVAH stand inventory, analysis and prescription program.

Marquis, D.A., Ernst, R.L., and Stout, S.L., 1984:Prescribing silvicultural treatments in hardwood stands of the Alleghenies. USDA Forest Service, Northeastern Forest Experiment Station, General Technical Report NE-96, 90 p.

Marquis, D.A., Grisez, T.J., Bjorkbom, J.C., and Roach, B.A., 1975:Interim guide to regeneration of Allegheny hardwoods. USDA Forest Service, Northeastern Forest Experiment Station, General Technical Report NE-19, 14 p.

Mendel, J.J., Grisez, T.J., and Trimble, G.R., Jr., 1973:The rate of value increase for sugar maple. USDA Forest Service, Northeastern Forest Experiment Station, Research Paper NE-250, 19 p.

Sander, I.L., Johnson, P.S., and Watt, R.F., 1976:A guide for evaluating the adequacy of oak advance reproduction. USDA Forest Service, North Central Forest Experiment Station, General Technical Report NC-23, 7 p.

Stout, S.L., and Nyland, R.D., 1986:Role of species composition in relative density measurement in Allegheny hardwoods. Canadian Journal of Forest Research, vol. 16, pp. 574-579.

Funding, Pests and Hardwood Log Grades: User-driven Cooperative Projects in the North Central States

M. R. Carroll
Forest Resource Assessment Supervisor
Forestry Division, Minnesota Dept. of Natural Resources
500 Lafayette Road, St. Paul 55155-4044

ABSTRACT

Users of Forest Inventory and Analysis (FIA) products require information that is both timely and tailored to local needs. In the North Central region, state agencies have actively promoted appropriation of federal funds to achieve a 10-year forest inventory remeasurement cycle. At the same time the states have made funds available to support increased survey intensity and newer FIA features of special interest such as habitat classification and fixed-radius growth plots. Other federal-state cooperative efforts are under way in the region. State and federal insect and disease specialists are working to standardize forest damage reporting; another program addresses the need for improved hardwood grade information.

ORGANIZATION AND OBJECTIVES

In March 1984 a Forest Inventory and Analysis Conference was held in Minnesota to heighten interest and increase participation among all North Central states. An advisory group was formed, and the following objectives were set:

- Establish recognition of the importance of FIA to states in the North Central Region, while at the same time affirming its importance to nationwide resource planning.
- Develop and adopt technologies and tactics that increase the usefulness and effectiveness of FIA, in order to minimize costs, assure adequate funding, and shorten the time period between cycles.
- Continue development of methods to make FIA data more accessible and usable to state forest agencies and other interested groups.
- Recognize and pursue ways to minimize social conflict while capturing more of the benefits available from forest land.

Since that time the North Central States FIA Working Group has cooperated with the North Central Forest Experiment Station of the U.S. Forest Service in monitoring the budgetary aspects of FIA in the region, reviewing survey costs, refining survey procedures and developing user-friendly access to the information. The participating states are encouraged to develop proposals and position papers reflecting the diversity of forest management concerns across the region; the site of the group's annual meeting is rotated from year to year.

MAJOR PROJECTS

One key issue has been the securing of adequate federal funds to assure a minimum FIA remeasurement cycle of 10 years at an intensity sufficient to provide plus or minus 3 per cent accuracy per million acres. State activity has centered on mailings and direct contacts with state foresters, governors and congressional delegations, emphasizing the need to restore and maintain adequate funding for North Central FIA. Federal appropriations have increased since 1984, and the Forest Service has redirected internal dollars in support of

FIA. In February of 1989 the Forest Service released new FIA schedules targeting a 10-year remeasurement cycle in the North Central States.

Special data collection above and beyond standard FIA specifications is the financial responsibility of the requesting state. Memoranda of Understanding are signed between the states and the Experiment Station for collection and processing of additional data. Apart from increases in sampling intensity, which several states have chosen to pursue, some users have initiated special studies: these include Wisconsin's habitat classification project, Michigan's private landowner survey, and Minnesota's fixed-radius growth plot study. These have been funded both through legislative appropriations and through contributions from client groups.

A project to standardize regional forest damange reporting brought together inventory and pest management specialists from all the states and from the Experiment Station. Forest Service State and Private Forestry funds went to Minnesota's Forest Pest Management Group, which worked with cooperators from all states to develop a standard list of categories of forest injury, a standard set of identification criteria, a set of diagnostic sheets for field use, and a slide-illustrated training package. The guiding aim is to place the analysis of forest pest losses on a comparable basis from state to state across the region. The system is to be tested in Minnesota and made available to all North Central states.

In recent years, technological change in hardwood-using industries coupled with over-harvesting of oaks in some North Central areas has left state utilization and marketing staff unsure of the volume and quality of hardwoods available for industrial use. Since 1985 the FIA Working Group has been testing new approaches to hardwood quality assessment. Missouri has applied the concept of "limiting factors," in which not only the log grade of a tree is recorded, but also the factor--diameter, form, or length--that keeps it from a higher grade. Starting in Iowa in 1989, federal FIA crews switched from log grading to tree grading. Representatives from Michigan, Wisconsin and Minnesota inventory units have attended training on the tree grading system, and are working with the North Central Station on ways to help clients understand and use the new data. Minnesota plans to test the grading of additional tree faces to help distinguish the better-quality Grade 3 trees. Efforts are also under way to improve recording of trees that do not meet grade specifications for clear faces but that may be important for local use and in construction.

The North Central FIA Working Group will continue to review the effectiveness of FIA procedures, promote adequate funding for the Station, and initiate refinements in development and delivery of FIA information. The things Experiment Stations can undertake on their own are limited by federal law and Forest Service policy, yet states require a diverse range of timely information to support their policies of economic development, multiple-use management, and environmental protection. User-driven cooperative projects are the answer.

Section 4

On the Estimation Process In Forest Inventory

Moderator: Carl Newton

Tiberius Cunia, Forest Inventory: On the Structure of Error of Estimates

Michael L. Clutter, Use of Regression Techniques in Industrial Forest Inventories

Edwin J. Green and William E. Strawderman, A Multivariate Composite Estimator

Francis A. Roesch, Jr., The Variances of Some Compatible Estimators for the Components of Growth From Remeasured Point Samples

R.B. Chevrou, The Truncated Law

Krishna P. Rustagi, Sampling Mixed Stands Using Product Estimator

Forest Inventory: On the Structure of Error of Estimates

Tiberius Cunia

Professor of Statistics and Operations Research
State University of New York College of Environmental Science and Forestry
Syracuse NY 13210

ABSTRACT

Because they are determined by sampling, the forest inventory estimates have error. It is customary to quantify this error in probabilistic terms, but because some of its sources are ignored, the error is generally underestimated. The paper analyzes the structure of the error by identifying its components (together with its sources) and by making some inferences about the size of these components.

INTRODUCTION

The forest biomass (volume) inventory is normally taken by statistical sampling. If the sampling method is properly designed and applied, and the resulting data are properly analyzed and interpreted, the inventory would normally lead to (i) estimates of the parameters of interest, (ii) significance tests of hypotheses about these parameters and (iii) inferences about the error of the estimates and the tests above.

For example, we may be talking about estimates of parameters such as mean biomass per acre, mean biomass growth per acre per year, etc. Or we may wish to test hypotheses such as mean biomass per acre of a 15 year old plantation is larger than 1000 cubic feet per acre, mean biomass growth has declined in the last five years (as compared to the growth of the previous five years), etc. Or, finally we may make inferences about error in the form of 95 percent confidence intervals for the mean biomass per acre or probability that we are wrong when we state that the mean biomass growth declined in the last five years, etc.

Great progress has been made during the last few decades in the methodology of selecting sampling units (plots or trees) and that of estimating parameters of interest. Very sophisticated sampling designs have greatly increased the efficiency of sampling by taking advantage of important theoretical developments in the statistical methodology and tremendous technological advances in computer hardware, remote sensing, etc. Little progress has been made, however, in one particular area; that of recognizing that, when the error of estimation is calculated, important error components are unaccounted for. As a direct consequence, the reports issued after the completion of a forest inventory may contain valid estimates of biomass but gross underestimates of their error; not all sources of error are well recognized with the corresponding error components evaluated and taken into account.

The objective of the present discussion is to give the reader an insight into the structure of the error of forest inventory. Because the way one looks at the error is subjective, so is the structure itself that one defines. The type of structure we shall define here was found useful in our studies. By analyzing this type of structure, one can (i) identify each error component and its source, (ii) evaluate the relative importance of each component (relative to the size of the total error), and (iii) obtain the means of controlling the size of the error components for a more efficient sampling design. In what follows, we shall borrow heavily on past papers, in particular borrow from Cunia (1987b). It is my hope that after this discussion, one may develop a "healthy skepticism" about the error statements contained in most reports on biomass estimates, by government or industrial forest organizations.

GENERAL STRUCTURE OF THE ERROR

In general, the error of forest biomass estimates has three main components. There is first the sampling error; the same sampling procedure applied repeatedly to the same forest population leads generally to selection of different sample units and, thus, to different estimates. And then there is the measurement error when the same sample units (trees or plots) measured by different people lead to different recorded values and, thus, to different estimates. Finally, the third error component is that of the statistical model used in deriving estimates; same inventory data analyzed and interpreted by different statisticians may lead to different estimates.

Each component in turn may consist of two main parts; a random error which tends to cancel itself out, in the long run with large samples and a systematic error or bias which is little affected, it at all, by the sample size. In the present discussion we shall not specifically deal with these two sub-components separately. We shall implicitly assume that the bias is zero or negligibly small, as we shall express the size of the error by its variance. Occasionally, when it is important we shall make reference to bias as part of the structure of error.

In some cases, the sampling design for forest inventory consists of one phase; the sample units (plots or trees) are selected from the forest area of interest and these units are measured for the variable of interest (biomass). Then, the error has the structure above with its three main components, the sampling, the measurement and the statistical model error. For example, this is the case with a 3P type of sampling in timber sales inventory, where the sample trees selected with unequal probabilities are measured for the variable of interest, the volume or value. Then, the sampling error is associated with the specific selection of sample trees, the measurement error with the difference between the recorded (as measured) and the true sample tree volume, and the statistical error with the bias of the 3P mean volume per tree, or total volume estimator.

It is extremely inefficient to apply one phase sampling design to the inventory of the biomass (say oven-dry weight) which may require destructive sampling. Then, the usual procedure is to use a two-phase design. In the first phase the trees of a relatively large sample are measured for "auxiliary" variables such as diameter, species, height, etc. which by themselves may not be of primary interest, but which are easy to measure and highly correlated with biomass, the variable of interest, which is not being measured. In the second phase, the trees of a small sample are measured for biomass in addition to some of the auxiliary variables of the first phase. The sample data of the second phase are used to determine a relationship (generally biomass regression functions) which applied to the data of the first phase sample yield the biomass estimates of the forest area being inventoried.

The error of the forest biomass estimates can still be viewed as consisting of sampling, measurement and statistical model components, with each component itself having two parts, one part associated with each of the two phases. We prefer, however, viewing the error as having two main components. The first one due to the sample of the first phase will be known here at the "main sample error" component and the second one due to the sample of the second phase will be known as the "biomass regressions error" component. Each of these two in turn will be viewed as having the sampling, the measurement and the statistical model error subcomponents.

It is inconvenient to have a relatively expensive and time consuming second phase sample, everytime one is contemplating taking the inventory of a given forest area. It is much better to construct biomass regressions for an entire region first, and then apply these regressions repeatedly to first phase types of samples, every time an inventory of a forest area within that region is being taken.

One may look upon these forest inventories as being taken by one-phase sampling designs, where the error of the biomass regression is considered as a part of the measurement error. Instead of actually measuring the tree biomass, one records the values as provided by the biomass regressions. The difference between the regression values and the actual tree biomass of the sample trees is the measurement error which may have a random part (due to the measurement of the auxiliary variables) and a systematic part (due to the difference between the expected values of the regression and actual tree biomass).

We prefer considering these forest inventory systems as being of a two-phase type. However, we must add another subcomponent to the main component error due to biomass regressions, that of application. The biomass regression functions were constructed for a tree population other than the one to which they are currently applied. The same biomass regression applied to different forest inventories will generate errors of different sizes.

Using the error structure above, with the two main components and each component consisting of three or four subcomponents, we can approach the problem of evaluating the error of the biomass estimates in its entirety, by considering all components and their sources. We consider first the fact that the error has two main components. To take them both into account one must find ways to appropriately express (i) the error due to the main sample (where the sample units are measured for the auxiliary variables alone), (ii) the error due to the biomass regressions, and then (iii) find a method to combine these two error components.

In what follows we shall discuss in more detail the structure of the main sample and of the biomass regression error components. Because the appropriateness of the error expressions is highly dependent on the method used to combine these components, we shall say a few words first about the way we have used to combine them.

COMBINING THE TWO MAIN ERROR COMPONENTS

More details about the methods of combining the main error components can be found, among others, in Cunia (1987a). In the short summary below, we shall use the variance as the expression of the error; it is more convenient to work with than standard error, mean square or confidence intervals, and its components are, at least in our case, additive. The disadvantage is that the bias is not included in the error and, thus, it is implicitly ignored.

To calculate a specific biomass estimate w, one must combine the statistics, say z of the first phase with the statistics, say b of the second phase sample. We shall assume that (i) $w = f(z,b)$ is an estimate of the parameter of interest $m = f(E(z), E(b))$ where E means mathematical expectation and (ii) estimates of the variances and covariances of z and b are available. When the first and second phase samples are statistically independent, the covariances between the z and the b statistics are all equal to zero. It can be shown then that the variance of w can be approximated by a long expression consisting of a sum of terms. Each term is a product of statistics z, b, their variances and covariances and values of partial derivatives of f with respect to z and b; the higher the order of differentials, the better the approximation.

The critical points of this approach are those of (i) expressing w as a simple function of z and b statistics and (ii) finding valid estimators of their variances and covariances. When the biomass regressions are non-linear, the expressions for w and its variance may be very complex or may not even exist in a closed form. But when the regressions are linear the expressions simplify tremendously. For this last case Cunia (1987a) presents the following approach.

The parameter m of interest is defined first as the product of two vectors $[m_z]$ and $[\beta]$, that is

$$m = [\beta]'[m_z] = \beta_1 m_{z1} + \beta_2 m_{z2} + ... + \beta_m m_{zm}$$

where $[\beta]$ is the vector of coefficients of the true regression function of the biomass y on some auxiliary variables $x_1, x_2, ..., x_m$ and $[m_z]$ is the vector of the expected values of some variables $z_1, z_2, ..., z_m$, functions of the auxiliary variables x. The estimators [z] of $[m_z]$ and [b] of $[\beta]$ are then defined from the first and second phase sample respectively, as well as estimators $[S_{zz}]$ and $[S_{bb}]$ of their covariance matrices. Finally, the statistics $w = [b]'[z]$ and $S_{ww} = [b]'[S_{zz}][b] + [z]'[S_{bb}][z]$ are defined as the estimators of m and variance of w respectively.

Note that the two components of the formula of S_{ww} may be interpreted as the error components of the first and second phase respectively. When the biomass regression has no error (or assumed to be without error), $[S_{bb}] = [0]$ and the variance of w reduces to the first component. On the other hand when $[S_{zz}] = [0]$, the case where the auxiliary variables of the entire population are measured without error, the first error component vanishes and the variance of w reduces to the second component. Finally, it can be shown that if the error components are expressed in relative, rather than absolute terms, the square of the percent standard error of w is equal to the sum of the squares of the percent standard errors of the two components. Stated otherwise, the percent error due to the first and second phase samples are additive in their squares.

ERROR COMPONENT DUE TO THE MAIN SAMPLE

The error component due to the main sample has been expressed as the covariance matrix $[S_{zz}]$. The way variances and covariances are usually calculated from the sample data would automatically take into account the random part of both, sampling and measurement error. As a matter of fact it would be extremely hard to separate them. The knowledge of their relative size is, however, only useful when one wishes to increase the efficiency of the sampling design. Should one measure with high precision a smaller number of sample units, or should one spend less time on their measurement and, thus, be able to measure more units?

Because the bias of sampling and measurement is not included in $[S_{zz}]$, it is important to insure that this bias is small relative to the random error. It may become important, even if small, when the sample size is very large. If the inherent bias in most sampling and measurement processes is to be minimized, the method of sample unit selection and measurement should be well defined and carefully applied. It may be interesting to note, however, that the effect of the measurement bias in biomass is zero, whenever, the same bias appears in the measurement of the auxiliary variables in both, the first and the second phase.

The covariance matrix $[S_{zz}]$ is calculated under the basic assumptions of a statistical model. Consequently, the error of the model is never accounted for. Empirical evidence seems to suggest that if the statistical model is selected with care, this error component is small. Various good models would ordinarily yield close results. This is the usual case when the estimation is done under the supervision of a good statistician. However, the error may be quite large when the selected model is poor. For example, applying simple random sampling formulae to the trees of randomly selected plots, would yield a gross underestimate of the error of the mean biomass per tree.

The size of the error component due to the main sample is a function of several factors. A look at the error component $[b]'[S_{zz}][b]$ shows that it is affected by (i) the sampling frame, sampling method and sample size (since the statistics z are generally averages of sample data), (ii) the parameter being estimated (different parameters are estimated by different statistics with different errors), (iii) the estimation procedure used (two estimators w_1 and w_2 of the same parameter may have different errors) and (iv) the auxiliary variables of the regression function (which determine the nature of the statistics z).

Research done at the College in the last five years has resulted in statistical methodology to derive the statistics z and their covariance matrix $[S_{zz}]$ for a variety of sampling methods and parameters to estimate. We have considered the basic designs of simple random sampling, stratified random sampling, cluster random sampling, two-stage sampling, double sampling for stratification, double sampling with regression estimators and sampling with partial replacement on two occasions. We have also considered estimates of the mean biomass per tree, mean biomass per acre, and mean growth per acre for various growth components (mortality, ingrowth, net change). We have finally considered several estimation procedures as functions of the sampling design and the parameter of interest. In all this methodology we have assumed linear biomass regression functions with estimates [b] and $[S_{bb}]$ given.

The results of this research are given in detail by Wharton and Cunia (1987). Applied to an actual set of data from 118 permanent sample plots and a biomass regression function calculated from 353 trees, the methodology showed a variety of results. The error component due to the main sample in percent of the total error was (i) about 80 percent for the estimation of the mean biomass per acre and mean biomass per tree, (ii) about 98 percent for the estimate of the net change in biomass between two successive measurements, (iii) about 40 percent for the estimates of the biomass growth on survivor trees, (iv) about 12 percent for the estimate of mean biomass per acre of ingrowth trees. Note that all these percentages refer to a relatively large sample of trees and a small sample of plots. Going from a sample of 118 to a sample of 472 plots (four times as large) and from a sample of 353 to a sample of 88 trees (four times as small) we have found that the error component due to the main sample represents only 20 percent of the total error of the mean biomass per acre. This implies that the corresponding percentage for the error of the biomass regression was 80 percent.

ERROR COMPONENT DUE TO BIOMASS REGRESSIONS

As stated before, the error of the biomass regressions has four main subcomponents. Let us discuss them separately in more detail.

Subcomponent 1 - Sampling Error

Defined as the error associated with the process of selecting sample trees, the sampling error is due to the inherent variation between the true biomass of the population trees. As with the first phase sampling error, the sampling frame and design, the sample size, the parameter to estimate and the estimator used will all affect the size of the sampling error.

Any sample of trees can be used to calculate a biomass regression. But for the resulting regression function to be statistically valid, the sampling design must satisfy three main conditions. First, the sample tree selection procedure must lead to a representative sample of the entire population of interest. This condition would be satisfied if every tree has a known (at least in relative terms), non-zero probability of being included in the sample. The second condition is that, for the resulting sample data, one must be able to devise a statistically sound method of estimation. What good would it do to have a representative sample if one could not analyze it properly? Finally, the selection procedure must be cost efficient; it must provide a sufficiently precise estimate of the regression function for reasonable sampling costs.

In an old survey, Cunia (1979 a,b) has found that most of the sampling and estimation procedures used to construct biomass regressions left much to be desired. For example, trees of "abnormal" shape were eliminated from the sample, even though the regression was expected to be applied to such trees as well. Sample trees may have been selected by procedures leading to representative samples but these procedures were so complex that they defied any attempt to properly analyze the data. Finally, to construct biomass regressions, regression techniques were used that were not consistent with the selection procedure. The sample was still representative but the standard least or weighted least squares methods used were not adequate; they had to be modified first.

The simulation research done at the College in the last few years and summarized in Wharton and Cunia (1987), shows the type of error one can make when, to representative samples of trees, selected by a variety of procedures, one applies the standard least and weighted least squares method. In some instances (as, for example, with stratified samples) the estimated biomass regressions were biased and most of the time the error of the regressions was grossly underestimated. This research strongly suggests that one should be extremely careful in the selection of the sampling and estimation procedures, when biomass regressions are to be constructed.

Subcomponent 2 - Measurement Error

Once selected, the tree must be measured for its biomass. It is useful to think of each tree as having a biomass which, for a given point in time, has a fixed value that is independent of the way the tree may be measured. For lack of a better name we shall denote this value as the conceptual or true biomass value of the tree. For practical considerations, however, one must work with values determined by a physical process of measurement. This process may require the complete measurement of the entire tree component biomass, or may be based on sampling. For example, the green weight of a given tree stem may be determined by the direct weighing of the entire stem, while the corresponding ovendry weight is usually estimated by subsampling.

It is convenient to assume that the measurement process can be applied (to a given tree) infinitely many times. Each time it is applied, a different biomass value may result. For lack of a better name, the resulting value will be known here as the recorded biomass value for the given tree. As the recorded values are random variables, it is useful to hypothesize the existence of the average of all possible recorded values, equal or unequal, an average known here as the expected biomass value of a given tree.

Note that we have first the true, conceptual biomass value which is defined independently of any measurement process, a theoretical value which probably exists only in our imagination. We then have the expected biomass value defined as the hypothetical average of all results of a well defined physical process of measurement, again a theoretical value. But because it is defined as the average of results that have a physical meaning, this definition has a much stronger inituitive appeal. Finally, we have the recorded biomass value, the result of a specific application of the measurement process. This is the only value we have and is the value used in all statistical calculations.

Because the expected value is defined in terms of a measurement process as it is actually applied in the real world, and this process implicitly assumes a computational procedure resulting in a recorded biomass value, this recorded value is theoretically an unbiased estimate of the expected value. For the case where the measurement process is defined in terms of subsampling, one may define, in addition, a value that would be obtained when the sample size becomes equal to the tree (or tree component) being measured; that is, when the sample is the entire tree (or tree component). Then, it is possible that the value we find by the complete biomass measurement of the tree may not be equal to the expected value of the recorded biomass. This may require a new definition, which, for lack of a better name we shall call the expected value of the complete measurement of the tree biomass.

The measurement error will be defined here as the difference between the recorded and the conceptual biomass value of a given tree. The measurement bias will be defined as the difference between the conceptual and the expected biomass value. The random measurement error will be defined as the difference between the recorded and the expected biomass value.

When the least or the weighted least squares method is used to estimate the biomass regression function and its error, the random subcomponent of the measurement error is automatically taken into account. Consequently, when a measurement process has a negligibly small bias (whatever definition of the bias one may use) the only question that may arise is that of whether one should measure intensively the biomass of a small number of sample trees (and, thus, reduce the measurement error within the sample trees at the expense of the sampling error between the trees) or measure with poor precision the biomass of a large number of sample trees (so that the sampling error between trees is reduced as much as possible).

Much theoretical and applied research is needed in order to find ways to (i) subsample trees for their biomass content by cost-efficient measurement and sampling techniques, and (ii) optimize sample size combination of number of sample trees and number of subsample units within the trees to either obtain desired precision for minimum sampling costs or maximize this precision for allowable costs of sampling.

The size of the measurement error is affected by the precision of the measurement process and by how much the expected biomass value differs from the true, conceptual biomass. When the measurement process is based on subsampling, the size of the error may also be affected by the precision of the subsampling technique, the efficiency of the estimator used and the difference, if any, between the expected value of the estimator and the expected value of the complete measurement.

In research done at the College we have analyzed subsampling procedures whereby (i) small wood samples are selected from the sample trees by random, stratified or cluster sampling and (ii) the tree biomass is estimated by simple or ratio type estimators. For these subsampling procedures we have derived the formulae for (i) the estimator of the tree biomass and its error and (ii) the optimal allocation of the financial resources between the number of sample trees and the size of the subsample within the sample trees. For more details on these techniques the reader is referred to two papers published in the Wharton and Cunia (1987) proceedings.

Sub component 3 - Error of Statistical Model

The error of the statistical model is due to the basic assumptions of the model which are seldom, if ever, satisfied by the real world. These assumptions may refer to the population of trees from which the sample trees are selected or to the sampling procedure itself. For example, some of the model assumptions may be that, for the population of interest, the true biomass regression function is of a specific form. Or, for a given sample of trees, the selection of trees has been made by the simple random sampling method. A simple thought suffices to show that, for a given tree population, the mathematical form of the regression function (if such a mathematical form exists) is unique and may never be known. Or, that the method of sample tree selection is seldom the simple random sampling; it is more efficient to draw trees in clusters or with probability proportional to some measure of tree size. Note that the error of the statistical model should not be confused with the error of specific estimates calculated under the model assumptions. This last error is usually an estimate of the sum of the sampling error and the random subcomponent of the measurement error and is as good as the model under which it is estimated.

The common procedure used in the calculation of the biomass regression functions is the least or weighted least squares as implicitly set by the standard computer software. The only decisions that one may have to make are those of the selection of (i) the set of independent variables and form of the regression function to use and sometimes (ii) the way the conditional variance is assumed to behave when it is not homogeneous. It is common to ignore the method by which the sample trees are selected; it is implicitly assumed that the selection is made by simple random sampling. Together with other assumptions, as for example the assumptions of normal distribution and form of regression function, this may be the main source of error of statistical model.

In the simulation research done at the College we have studied the effect of both sampling design and estimation procedure on three sets of statistics: (i) the point estimate, (ii) the error (bias and variance) and (iii) the sample-based estimate of the error of the sample average biomass per acre. The sampling procedures that we have used in our simulation process have been briefly described above. The estimation procedures that we have applied to samples selected by various methods can be classified into two main groups; the first group based on the ordinary least and weighted least squares method as used in the standard computer packages and the second group based on the least and weighted least squares modified to take into account the sample tree selection procedure. We have varied, in addition, the form of the regression function and the set of independent variables included in these regressions. For more details on the sampling and estimation procedures the reader is referred to the Wharton and Cunia (1987) proceedings.

The results obtained are as follows. If the form of the regression function is properly selected, the effect of the statistical model on the point estimate of the average biomass per acre is negligibly small. With the possible exception of stratified sampling, this effect is largely independent of the sampling method or the estimation procedure used. On the other hand, the effect of the statistical model on the estimation of the error or the associated interval estimates of the average biomass per acre or hypothesis testing, may be considerable. In particular, the use of the ordinary least or weighted least squares method to estimate the error of the biomass regression and to calculate confidence intervals or test null hypotheses is highly questionable. Consequently, all inferences of error made on the basis of statistics obtained from the application of the standard statistical computer packages should be viewed with suspicion and a heavy dose of skepticism is highly recommended.

The bias seems to be affected by the probability of tree selection. It is practically zero for equal selection probability, it is slightly different from zero (but still not significantly different from zero) when the sampling is done with probability proportional to tree height h or tree diameter d and it becomes much larger and significantly different from zero when the probability of selection becomes proportional to basal area d^2 or approximate volume d^2h. This is a surprising result. However, before accepting this conclusion as true, further research is highly recommended. For now, one should keep an open mind and consider the possibility of obtaining biased results, when the tree selection is made proportional to basal area or volume. The bias may also be critically affected by the allocation of the sample units among strata, when stratified sampling is used and the effect of stratification is not properly accounted for by the estimation procedure. For the particular case of stratified sampling with proportional allocation the bias is expected to be small, if any. If the allocation is not proportional to stratum size, the indiscriminate use of the ordinary least or weighted least squares procedures may lead to a bias which is generally large and significantly different from zero.

It came as a complete surprise to find out that (i) the precision of the estimates is only slightly improved when the weighted least squares are used instead of the least squares; the least squares techniques are much more robust than we ever suspected, and (ii) the precision is not necessarily improved when the sample trees are selected with probability proportional to a measure of tree size; there is even evidence suggesting that the precision of the estimators decreases as the probability of selection increases with the tree size. This last conclusion goes so much against our intuition that further research must be done before accepting it as true.

The only known estimate of the precision of the regression function is the estimate derived from the analysis of a given sample under the given statistical model. The simulation results show that the precision is grossly overestimated (the error is grossly underestimated) by the ordinary least and weighted least squares method. On the other hand, the precision seems to be slightly underestimated (the error is slightly overestimated) by the modified least or weighted least squares that we have used. This last result may be due to the fact that we have ignored the effect of the finite population correction factor when calculating the estimate of the error, or due to the sampling error, or simply due to the fact that the modified procedures do indeed overestimate the error.

Subcomponent 4 - Application Error

Defined as the part of the error associated with a specific use of the biomass regression, the error of application is due to the difference between the true biomass regression functions of (i) the population of trees for which the biomass regressions were derived and (ii) the population to which these regressions are currently applied. Seldom, if ever, these two populations have identical regression functions, no matter how similar they are thought to be. The populations are dynamic in nature and change continuously with time. By the time the biomass regression is estimated, the population from which the sample trees were selected has already changed to a new population with new characteristics. Also, for practical purposes, it is common to construct biomass regression functions for one tree population and then apply them repeatedly to many similar populations at future points in time. Finally, the same biomass regressions applied to parts only of the population of trees, or applied to estimation of different parameters may lead to errors of different size. For example, the error of the estimate of the ingrowth biomass per acre is generally much larger in size than the error of the estimate of the overall average biomass per acre, even though the same regression function is being applied in each case. For more on the case where regional biomass regressions constructed for the all size trees of a given population are applied to the trees of a specific size only, the reader is referred to Yang and Cunia (1989).

It does not seem that appropriate statistical theory exists to take into account the error of application. However, it is conceptually possible to use basic statistical theory and derive specific methodology to quantify the error of application of biomass regressions by subdividing the "total" error of the biomass regressions into subcomponents associated with (i) differences between geographical regions withing the large forest area for which the regressions are calculated, (ii) differences between forest stands within regions, (iii) differences between classes of trees (ingrowth, survivors, dominant trees, etc.) within stands, and so forth. One way to do this could be that of using dummy variables to include in the regression function geographical regions, forest stands, classes of trees, etc.

FINAL REMARKS

It is common to report forest inventory estimates with their error, and it is also common to assume that this error is realistically evaluated. Because all the sources of the error may not be properly accounted for, chances are that this error is grossly underestimated.

In the present paper, we have analyzed the structure of the error of estimation and identified three main error components: the sampling, the measurement and the statistical model error. These components appear in the error of both, the main sample of plots and the biomass regressions. When regional or standard biomass regressions are being used, there is a fourth error component, that associated with the application of the regressions.

Some of the error components may be small; like that of statistical model, when the model is well selected, or that of measurement, when carefully made. Others, however, like that of the sampling error and application of the biomass regressions may become quite large. Because these components are seldom, if ever accounted for, one should view all statements of error of estimation with great care.

REFERENCES

Cunia, T., 1979a: On tree biomass tables and regression: some statistical comments. In: W.E. Frayer, Ed., Forest Resources Inventory, Proceedings of a Workshop, Colorado State University, Department of Forest and Wood Sciences, Fort Collins, CO.

Cunia, T. 1979b; On sampl.ing trees for biomass tables construction: some statistical comments. In: W.E. Frayer, Ed., Forest Resources Inventories, Proceedings of a Workshop, Colorado State University, Department of Forest and Wood Sciences, Fort Collins, CO.

Cunia, T., 1987a: Error of forest inventory estimates: its main components. In; Wharton, H.E. and Cunia, T. Eds., Estimating the biomass regressions and their error, Publication Nr. NE-GTR-117, USDA Forest Service, Northeastern Forest Experiment Station, Broomall PA.

Cunia, T., 1987b: On the error of biomass regressions and the corresponding inventory estimates. In: Proceedings of the Ninth Annual Southern Forest Biomass Workshop, Biloxi, Mississippi, Mississippi State University.

Wharton, H.E. and Cunia, T., 1987: Estimating tree biomass regressions and their error, Publication Nr. NE-GTR-117, USDA Forest Service, Northeastern Forest Experiment Station, Broomall PA.

Yang, Xiaoming and Cunia, T., 1989: Construction and Application of biomass regressions to size classes of trees. In: Proceedings of the "State of the Art Methodology of Forest Inventory" July 30-August 5, 1989, Syracuse NY.

Use of Regression Techniques in Industrial Forest Inventories

Michael L. Clutter
Research Scientist, Forest Biometrics
Union Camp Corporation
P. O. Box 1391
Savannah, GA 31402

ABSTRACT

The development of regression techniques, particularly in forest growth and yield modeling, is discussed, and their use in southern industrial forest inventory is explored. Several examples are given which document the use and misuse of regression analysis in forest inventory.

Inventory design considerations allowing for proper interface with common growth and yield models are explored. In many instances, these design considerations suggest alternative inventory procedures for estimating forest characteristics of interest.

INTRODUCTION

During the past few decades, intensive forest management has increased significantly throughout the southeastern United States. To aid in the formulation of informed management decisions, inventories describing forest conditions were developed. Originally, the majority of these inventories were used to address long term raw materials supply issues and generally were not designed to provide individual stand-level management planning information. Many such systems developed as continuous forest inventories and, as such, provided reliable information on forest wide trends - but did little to address individual stand management decisions.

With the advent and application of quantitative asset valuation techniques, a need for stand level forest inventory information developed. Given this information, along with a growth and yield model, forest owners could quantitatively evaluate the effects of different management scenarios. As interest in preparing these "what-if" scenarios increased, the demands on both growth and yield modeling and forest inventory information expanded concordantly.

The recent history of forest inventory design and the development of growth and yield models in the southeast is closely tied to the application of regression techniques. Individual tree volume and weight equations, explicit yield prediction systems, and height versus age functions represent just a few of the early applications of regression analysis in forest inventory. In addition, sampling designs which utilize regression also become popular. Such methods as double sampling with a regression estimator, point-model based sampling (Schreuder, 1984), sampling with partial replacement for growth (Ware and Cunia, 1962) had regression techniques embedded within the sampling design calculations. Obviously, these are just a few of the applications of regression analyses used in forest inventory. Below is a discussion of three different situations in forest inventory where regression is extensively used: (1) regression techniques implicit to sampling designs, (2) regression techniques for calculating stem weight and volume, and (3) regression techniques for stand growth and yield modeling.

Regression Techniques Implicit to Sampling Designs.

Several commonly used sampling schemes employ regression techniques in their sampling design calculations. All of these designs attempt to increase precision by the use of an easily measured variate that is correlated with the variable of interest. Regression estimators are used in several industrial applications in the southeast.

An often utilized application in the south, particularly in procurement inventory situations, is the use of double sampling. This scheme takes advantage of the strong linear relationship between basal area and volume in forest stands. Usually, a large sample of prism points is installed in the stand of interest, and the number of "in" trees are counted to provide an estimate of basal area. On a random subsample of points, individual tree dbhs and heights are tallied so volume can be calculated using a regression estimator which relates basal area to volume. Due to the strong linear relationship, this type of sampling scheme consistently provides more efficient estimates than does standard prism cruising workup. One problem, however, which can arise is that sample sizes for the subsample can be smaller than desired when estimates by species and product class are required. Care must be taken to adequately sample each population of interest if such breakdowns are of interest. Experience has shown this technique to work best in young pine plantations which are almost exclusively one product - pine-pulpwood.

Other sampling designs which use regression estimators include point-model based sampling and sampling with partial replacement for growth - an application closely analygous to the one delineated above. No doubt other designs not addressed, likely exist which utilize regression estimators as an integral part of the sampling scheme.

Regression Techniques for Calculating Stem Weight and Volume.

An important use of regression analysis in forest inventory is the computation of individual stem volume and weight. For instance we currently employ the use of over 100 volume or weight equations for inventory analyses in the Woodlands Division at Union Camp. Whether the application be procurement oriented or management oriented, a volume equation applicable to the population of interest is required to calculate volumes. The problem can be formulated as:

$$Y = f(D,H,F)$$

where Y = some measure of tree content
D = a function of tree dbh
H = some measure of tree height
F = some expression of tree form

Most common applications of volume or weight prediction do not involve any explicitly measured function of tree form. Equations which do not have form as an independent variable are less expensive to use because no estimate of upper stem diameter is required. Many different functional forms have been utilized (see Clutter, et.al., 1983) for volume or weight equations.

One problem associated with the fitting of tree content equations is unequal variance over the range of the dependent variable. The assumption of constant variance for the dependent variable is rarely satisfied as required by standard regression assumptions. To correct for unequal variance problems, the model should be transformed by dividing all variables by a factor proportional to the standard deviation of the dependent variable. Such a transformation will produce a model which exhibits constant variance. As pointed out by Cunia (1964), the variance of stem content is proportional to either D^2H or $(D^2H)^2$ for most tree populations.

Inventory Designs on Fee Simple Lands

Most current industrial forest inventories on fee simple lands have evolved into roughly the following format. Stands are sampled, in most cases using point sampling techniques, on some type of systematic grid. The number of points or plots installed is generally a function of stand size and sampling intensities are almost always quite light. In most cases, all stands in a relatively contiguous unit are sampled so current stand age isn't considered a factor in when to sample a given stand. Each stand is usually remeasured every five to eight years, and those stands not measured are updated using an applicable growth and yield model developed with standard least square techniques. The minimum information related will include dominant and co-dominant height, stems per acre, stand age, basal area, a stand diameter distribution, and some type of soil management class information which may reflect differences in trafficability, stand productivity, and also aid in projecting growth and yield.

As mentioned above, those stands not cruised in a given year are updated using an applicable growth and yield model. Currently, two types of models are being utilized for these updates; either a diameter distribution model (for example see Dell, et.al., 1979) or an individual tree distance independent model similar to that reported by Clutter and Allison (1974) or Pienaar and Harrison (1988). Several problems occur when either of these types of growth and yield models are used to perform inventory updates.

Diameter distribution models have been the preferred method for inventory updates for the last fifteen years in southern pine plantations. But since diameter distribution information is almost always collected in the sampling process, the ability of the model to utilize this information as input is lacking - hence the information was never put to good use. Another problem which perplexed many users of inventory information was that of shrinking volumes for certain product classes - also a function of using a diameter distribution model. This problem occurs when the stand is sampled the distribution obtained from the sample is stored in the inventory file. When the stand is updated the following year, situations arise where the updated distribution shows less volume than the sampled distribution from the previous year. While in most cases the differences are well within the expected error for the growth and yield model, such inconsistencies are hard to explain to users of the system.

Because of the problems identified above, several companies have started using generalized stand table projection algorithms (see Harrison and Daniels, 1988) to update stands. These models use stand-level variables as input to a set of core equations such as these:

$$H_2 = \{H_1^{e^{\beta_1(1/A_1-1/A_2)}}\} \{\exp[\beta_2((1/A_2)-(1/A_1)e^{\beta_1(1/A_1-1/A_2)}) + \beta_3(e^{\beta_1(1/A_1-1/A_2)}-1)]\}$$

$$N_2 = [\, N_1^{\alpha_1} + (\alpha_2 + \alpha_3/H_2)(A_2^{\alpha_4} - A_1^{\alpha_4})]^{\;1/\alpha_1}$$

$$B_2 = B_1 \exp[\delta_1(1/A_2-1/A_1)\; N_2^{\delta_2+\delta_4(1/A_2)}\; N_1^{-\delta_2-\delta_4(1/A_1)}\; H_2^{\delta_3+\delta_5(1/A_2)}\; H_1^{-\delta_3-\delta_5(1/A_1)}$$

Where:

A_i = stand age at time i
H_i = dominant and co-dominant height at time i
N_i = stems per acre at time i
β_i = basal area per acre at time i
δ's, β's, σ's are parameters to be estimated

In addition, using the sampled diameter distribution as input alleviates the problems associated with diameter distribution models. One problem associated with this type of model is the lack of path invariance - which requires the user to retain the last cruise on the inventory file as a basis for future projections. We have been pleased with the behavior of such models, and they appear to perform more accurately over short projection periods of one to five years in length.

A central problem with projection models, however, is their inability to maintain consistent relationships between such variables as volume, quadratic mean diameter, and surviving stems when each model involved is fitted independently. Recent advances in fitting systems of simultaneous linear or nonlinear equations have essentially solved this problem (Borders, 1989). In addition, these techniques, along with biologically reasonable model forms, have ensured that such models are consistent with basic biological principles.

Another problem of significant concern in the application of growth and yield models for inventory updates is the issue of aggregation of plots within stands. Can we safely use stand averages of such quantities as basal area or stems per acre to perform inventory updates, or should the individual plots be projected and averages computed. Our experience has shown that this depends on the model in question. If, for example, the core equations presented above were appropriate then, due to the polymorphism present in the site index equation and the presence of a term reflecting site quality in the mortality function, averaging plots would not be advisable. If site index equation is anamorphic and no effect of site on mortality is present, using stand averages as input usually has little effect on projected yields. At times, if polymorphic trends are present, differences of thirty to forty percent in projected yields are not uncommon between the two methods.

SUMMARY

This paper is a review of the application of regression techniques in southern industrial forest inventory. Mentioned are only a few of the applications for which regression analysis is appropriate. In addition, several common problems were discussed which are an outcome of the application of this type of analysis. Significant problems still exist as to what are the proper methods to sample and project forest inventory information. Opportunities exist to find more efficient methods of utilizing data collected in the forest inventory sampling process.

REFERENCES

Borders, B.E., 1989: Systems of equations in forest resources modeling. For. Sci. In press.

Clutter, J.L. and Allison, B.J., 1974: A growth and yield model for pinus radiata in New Zealand. J. Fries, ed. Growth model for tree and stand simulations. Royal College of Forestry, Stockholm, Sweden. p. 136-160.

Clutter, J.L. and Fortson, J.C., Piennar, L.V., Brister, G.H., Bailey, R.L., 1983: Timber management: a quantitative approach. John Wiley & Sons. New York, NY.

Cunia, T., 1964: Weighted least squares method and construction of volume tables. For.Sci. 18:241-285.

Dell, T.R., Feduccia, D.P., Campbell, T.E., Mann, W.F., and Polmer, B.H., 1979: Yields of unthinned slash pine plantations on cutover sites in the west gulf region. USDA-FS Res. Pap. SO-147.

Pienaar, L.V. and Harrison, W.M., 1988: A stand table projection approach to yield prediction in unthinned even-aged stands. For.Sci. 34:804-808.

Schreuder, H.T., 1984: Poisson sampling and some alternatives in timber sales sampling. For. Ecol. Mgmt. 8:149-160.

Ware, K.D. and Cunia, T., 1962: Continuous forest inventory with partial replacement of samples. For. Sci. Monograph 3. 40 pp.

A Multivariate Composite Estimator

Edwin J. Green
Associate Professor of Forestry
Cook College — Rutgers University
P.O. Box 231
New Brunswick, NJ, 08903

William E. Strawderman
Professor of Statistics
Rutgers University
P.O. Box 231
New Brunswick, NJ, 08903.

ABSTRACT

The usual method of combining sample data with auxiliary information is the familiar precision—weighted composite estimator. However, application of this estimator is straightforward only when the auxiliary information is unbiased. If this is not the case, and the bias is unaccounted for, then the risk of the usual composite estimator can be greater than that of the sample mean. We conjecture that investigators are often unsure of the possible bias in their auxiliary information. Accordingly, we present an estimator which mimics the usual composite estimator when the auxiliary information is unbiased, yet dominates the sample mean even as the bias becomes large. In addition, unlike the usual composite estimator, the new estimator does not require one to specify the variance of the auxiliary information. The performance of both estimators, and the sample mean, is examined based on results of a simulation trial on actual sample data.

INTRODUCTION

Occasionally, a forester has more than one source of information with which to estimate a given parameter (e.g., cubic volume per unit area). For example, in addition to sample data drawn from the specific stand(s) of interest, he/she may have data from another similar stand or a predicted value for the parameter based on a model calibrated with data from other stands. If the size of the sample drawn from the stand of interest is adequate, the forester may be satisfied with an estimator based solely on these data. However, if the sample size is small, he/she may opt to use the auxiliary information in order to increase the precision of the estimate. Estimators which combine information from various sources are variously referred to as synthetic, combined, and/or composite estimators, among other terms. It is well known that if the component estimators are independent and unbiased scalars, then the optimal method of combining them (assuming quadratic loss) is to weight each component by the inverse of its variance. For example if there are two components, X and Y, the optimal composite estimator is:

$$\delta_{\mathrm{comb}} = \mathrm{c}\,\mathrm{X} + (1-\mathrm{c})\,\mathrm{Y} \tag{1}$$

$$\text{where: } \mathrm{c} = \mathrm{Var}(\mathrm{Y})/\{\mathrm{Var}(\mathrm{X}) + \mathrm{Var}(\mathrm{Y})\},$$

and Var(X) denotes the variance of X. If the respective variances are known then δ_{comb}

dominates the sample mean, say X. Should the auxiliary component estimator (Y) be biased then its variance should be replaced by its mean square error in (1). If this is not done and Y is mistakenly assumed to be unbiased, then c may be estimated poorly, resulting in an estimator with higher risk than X. We conjecture that the forester may often be unaware of or unsure of the bias of Y. Accordingly, we present an alternative composite estimator in which the bias of Y need not be specified. The new estimator mimics the risk of δ_{comb} when the bias of Y is small, yet improves upon X even as the bias of Y becomes large. The estimator is based on the James—Stein result (James and Stein 1961), and requires that one be interested in three or more problems. We present the estimator in the following section and then conclude with an example application of the estimator to real data.

DEVELOPMENT OF ESTIMATOR

Let **X** and **Y** be independent p—variate normal random vectors. Assume the covariance matrices are $\sigma^2\mathbf{I}$ and $\tau^2\mathbf{I}$ respectively, where **I** is the identity. Further assume (for the moment) that σ^2 and τ^2 are known. Let $E(\mathbf{X}) = \boldsymbol{\theta}$ and $E(\mathbf{Y}) = (\boldsymbol{\theta} + \boldsymbol{\xi})$ be unknown, where $\boldsymbol{\xi}$ is the bias of **Y**. The objective is to estimate $\boldsymbol{\theta}$ with minimum loss. We will primarily be concerned with the usual quadratic loss: $\|\delta - \boldsymbol{\theta}\|^2 = \Sigma\,(\delta_i - \theta_i)^2$, where δ is an estimator for $\boldsymbol{\theta}$.

In this paper we study an estimator which combines the unbiased estimator **X** and the biased estimator **Y** in the spirit of James—Stein estimation by shrinking **X** toward **Y**:

$$\delta_{p-2}(\mathbf{X},\mathbf{Y}) = \mathbf{Y} + \left(1 - \frac{(p-2)\sigma^2}{\|\mathbf{X}-\mathbf{Y}\|^2}\right)(\mathbf{X}-\mathbf{Y}) \tag{2}$$

$$= \mathbf{X} + \left(\frac{(p-2)\sigma^2}{\|\mathbf{X}-\mathbf{Y}\|^2}\right)(\mathbf{X}-\mathbf{Y})$$

It is shown in Green and Strawderman (1989) that the risk, or expected loss, of the estimator (2) is given by (for $p \geq 3$):

$$E\|\delta_{p-2}(\mathbf{X},\mathbf{Y}) - \boldsymbol{\theta}\|^2 = p\sigma^2 - (p-2)^2\sigma^4\, E\left(\frac{1}{\|\mathbf{X}-\mathbf{Y}\|^2}\right). \tag{3}$$

Note that the risk of the MLE **X** is $p\sigma^2$. Hence, since σ^2 and $\|\mathbf{X}-\mathbf{Y}\|^2$ are positive, δ_{p-2} dominates **X**.

Green and Strawderman (1989) show that it is only necessary for **X** to be normally distributed in order for $\delta_{p-2}(\mathbf{X},\mathbf{Y})$ to dominate **X**. Since **X** will commonly arise from a random sample in forestry applications, we may appeal to the Central Limit Theorem to justify the assumption that **X** is normally distributed.

Jensen's inequality (see, e.g., Lehmann 1983) applied to (3) gives an upper bound on the risk of $\delta_{p-2}(\mathbf{X},\mathbf{Y})$ which we present in the following:

$$E\|\delta_{p-2}(\mathbf{X},\mathbf{Y})-\boldsymbol{\theta}\|^2 \leq p\sigma^2 - (p-2)^2\sigma^4\,\frac{1}{E\|\mathbf{X}-\mathbf{Y}\|^2} = p\sigma^2 - \frac{\sigma^4(p-2)^2}{\boldsymbol{\xi}'\boldsymbol{\xi} + p(\sigma^2+\tau^2)}. \tag{4}$$

If **Y** is in fact unbiased, i.e., $\boldsymbol{\xi} = \mathbf{0}$, the bound in (9) becomes $p\sigma^2 - [(p-2)^2\sigma^4/p(\sigma^2+\tau^2)]$.

However, in this case the risk (3) may be evaluated exactly since $\|\mathbf{X}-\mathbf{Y}\|^2/(\sigma^2+\tau^2)$ is a central χ^2 with p degrees of freedom. It is known that $E(1/\chi_p^2) = (p-2)^{-1}$. Hence, for $\boldsymbol{\xi} = \mathbf{0}$:

$$\begin{aligned} E\|\delta_{p-2}(\mathbf{X},\mathbf{Y}) - \boldsymbol{\theta}\|^2 &= p\sigma^2 - \frac{(p-2)\sigma^4}{\sigma^2+\tau^2} \\ &= p\left[\frac{\sigma^2\tau^2}{\sigma^2+\tau^2}\right] + \left[\frac{2\sigma^4}{\sigma^2+\tau^2}\right] \end{aligned} \tag{5}$$

The multivariate version of the usual precision—weighted combined estimator is:

$$\delta_{\text{comb}}(\mathbf{X},\mathbf{Y}) = \left[\frac{\tau^2\mathbf{X}}{\sigma^2+\tau^2} + \frac{\sigma^2\mathbf{Y}}{\sigma^2+\tau^2}\right]. \tag{6}$$

It is easy to show that the risk of δ_{comb} (assuming both **X** and **Y** to be unbiased for θ) is $p\left[\frac{\sigma^2\tau^2}{\sigma^2+\tau^2}\right]$. As stated earlier, it is well known that δ_{comb} is the optimal component—by—component linear combination of **X** and **Y** under quadratic loss when both estimators are unbiased. Thus if both **X** and **Y** are unbiased the risk of δ_{p-2} is larger than that of δ_{comb} by $\left[\frac{2\sigma^4}{\sigma^2+\tau^2}\right]$, or equivalently by the proportion $\left[\frac{2\sigma^2}{p\tau^2}\right]$. Therefore, if σ^2 and τ^2 are of similar magnitude and p is large, the penalty for using δ_{p-2} instead of δ_{comb} is small.

In contrast, if **Y** is not unbiased for $\boldsymbol{\theta}$, i.e., $\boldsymbol{\xi} \neq \mathbf{0}$, the risk of the James—Stein combined estimator is bounded by $p\sigma^2$ while the risk of the usual composite estimator (if **Y** is mistakenly assumed to be unbiased) can be shown to be:
$\left[p\left[\frac{\sigma^2\tau^2}{\sigma^2+\tau^2}\right] + \left[\frac{\sigma^2}{\sigma^2+\tau^2}\right]^2\xi'\xi\right]$, which is unbounded.

We thus have a combined estimator δ_{p-2} which: 1. improves on the unbiased estimator **X**; 2. does nearly as well as the usual combined estimator when both **X** and **Y** are unbiased; and 3. does much better than the usual combined estimator when **Y** is badly biased.

Of course, if **Y** is biased, then δ_{comb} is not the proper precision weighted estimator, and τ^2 should be replaced by $(\tau^2+\xi'\xi)$ in (6). However, in a practical sense this means that one must be able to estimate the MSE of **Y**. This may be a problem, and it is not entirely clear what the optimal procedure is in such a case (Green and Strawderman 1988). However, it is not necessary to estimate the MSE of **Y** with δ_{p-2}. In fact, one does not even need to specify τ^2 in δ_{p-2} (in contrast to δ_{comb}, which depends on both σ^2 and τ^2).

In practice, σ^2 is usually unknown. In this case, if s^2 is an unbiased estimate of σ^2 independent of **X** and **Y** such that ks^2/σ^2 is distributed as χ^2 with k degrees of freedom, then $(p-2)\sigma^2$ should be replaced in δ_{p-2} by $s^2[k(p-2)/(k+2)]$.

The risk of the three estimators (**X**, δ_{comb}, and δ_{p-2}) is plotted against the bias (ξ) for the case $\sigma^2=1$, $\tau^2=1$, and $p=25$ Figure 1. In the Figure it is assumed that the bias of the Y_i is constant, i.e., $\xi_i = \xi \; \forall$ i. Among the features evident in the Figures are: i. δ_{comb} has the lowest risk of the three estimators depicted when $\xi = 0$, but the risk of δ_{comb} increases quadratically as ξ increases, becoming greater than the risk of **X** or δ_{p-2}; and ii. the risk of δ_{p-2} is close to the risk of δ_{comb} initially ($\xi=0$), but as ξ increases the risk of δ_{p-2} approaches but never exceeds the risk of **X**.

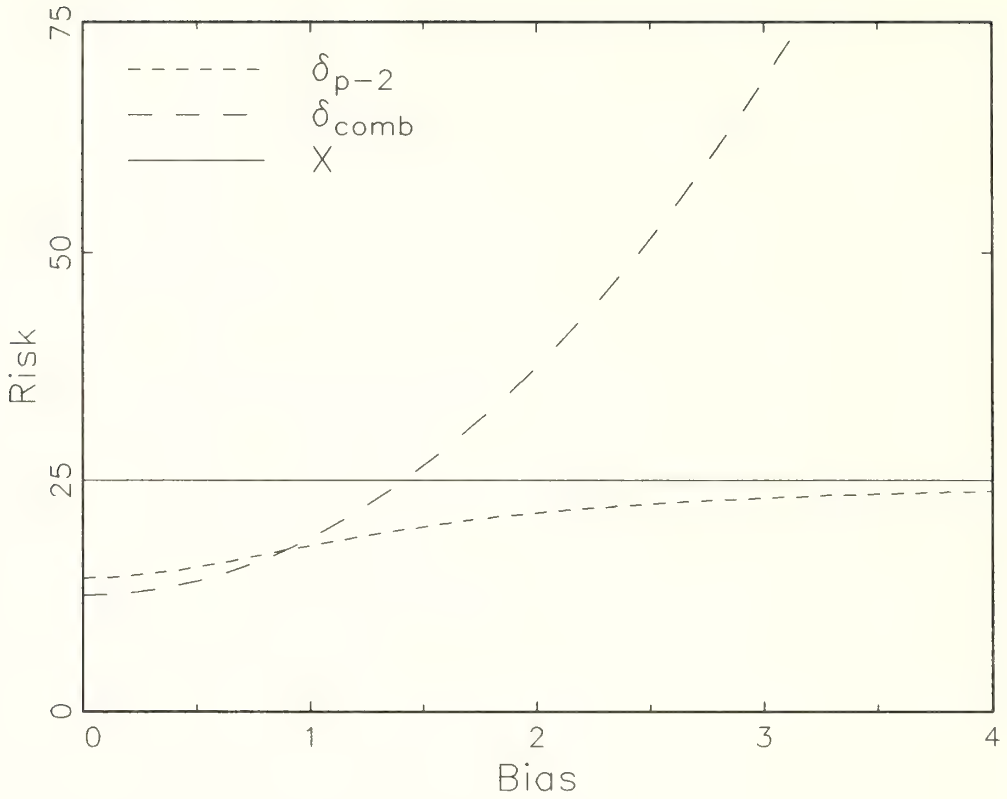


Figure 1. Risk vs. bias for X, δ_{comb}, and $\delta_{p\text{-}2}$. Here p=25, $\sigma^2=1$ and $\tau^2=1$. Constant bias across dimensions is assumed, i.e., $\xi_i=\xi\ \forall$ i.

One obvious improvement over $\delta_{p\text{-}2}$ is the positive part version

(7) $$\delta_{p\text{-}2}{}^{+}(\mathbf{X},\mathbf{Y}) = \mathbf{Y} + \left(1 - \frac{(p-2)\,\sigma^2}{\|\mathbf{X}-\mathbf{Y}\|^2}\right)_+ (\mathbf{X}-\mathbf{Y})$$

$$\text{where } (Z)_+ = \begin{cases} Z \text{ if } Z \geq 0, \\ 0 \text{ if } Z < 0. \end{cases}$$

It can be shown that $\delta_{p\text{-}2}{}^{+}$ dominates $\delta_{p\text{-}2}$.

APPLICATION

Suppose we are interested in estimating the basal area per acre (B) of a number (≥ 3) of loblolly pine stands. To this end, independent inventories are performed in each stand. Further suppose we have an independent model which predicts B given the age of the stand (A), number of surviving stems per acre (N), and the height of the dominants and codominants (H). It may be reasonable to suppose that the efficiency of the estimate of basal area per acre could be improved by combining the sample mean with the model prediction for that stand. However, as Green and Strawderman (1988) point out, model predictions tend to be biased. This occurs for at least two reasons: First, most models predict *ln*(B) and it is well known that such models are biased for B. Second, these models are typically fitted to many stands with numerous combinations of levels for A, N, and H. The stand effect is a random effect, even though it is not usually explicitly included in the model. Thus even if the model was unbiased for the mean B of *all* stands with a given level of A, N, and H, it would still in all likelihood be biased for a *particular* stand. The latter point was also noted by Lappi and Bailey (1988). Green and Strawderman (1987), using actual data from loblolly pine stands, report an instance where model bias apparently caused the usual composite estimator ($\delta_{\mathrm{comb}}(X,Y)$) to have a *higher* MSE than the sample mean.

Here we report the results of an experiment conducted to simulate the scenario described above. The important difference between our experiment and the envisioned scenario is that our data were such that none of the basal area yield models in the literature were appropriate. Thus we were forced to use a subset of the data to estimate the parameters of an arbitrarily chosen model. Therefore the model estimates and the sample means are not completely independent. This should favor the composite estimators over the sample mean. However, as only a small percentage of the data (<5%) were used for fitting the model, we believe that the violation of the independence assumption is of minor importance. Additionally, we did not screen a number of alternative models in order to find the one which fitted the data best. Instead, we simply employed a model form that has been used elsewhere (Parresol and others 1987). This was done to avoid biasing our results heavily in favor of the composite estimators by using the optimal model for our data set.

Data:

For this experiment we used data provided by three Southeastern forest products companies (Westvaco, Union Camp, and International Paper). The data consisted of 1126 sample plots in 66 loblolly pine stands. Company 1 contributed data from 653 plots in 47 stands, while companies 2 and 3 provided data from 143 plots in 9 stands and 330 plots in 10 stands, respectively. Observations on B, A, N, and H were made at each plot. The number of plots per stand varied from 5 to 50, with a mean of 17.

Simulation:

We simulated the performance of 3 estimators, $\bar{X}$, δ_{comb}, and $\delta_{p-2}{}^{+}$ as follows: First, the mean B per stand ($\bar{B}_i{}^*$, where * signifies that all the data from stand i were used) was computed. These values were assumed to be equal to the θ_i's. The simulation then proceeded in two stages. In the first stage, 50 plots were randomly drawn from the data using proportional allocation among the companies. The following model was fitted by ordinary least squares to these 50 observations:

$$ln(B) = \beta_0 + \beta_1 A^{-1} + \beta_2 ln(N) + \beta_3 ln(H) + \epsilon. \tag{8}$$

Next, five observations on B_{ij} (where the subscripts refer to stand and plot, respectively) were randomly selected (with replacement) from the data remaining for each stand, i.e., the 50 observations used to fit the model were excluded from this stage. The sample mean ($\bar{B}_i$) was calculated for each stand, as were δ_{comb} and δ_{p-2}^+. For the latter two estimators, the sample means were combined with model predictions.

It is well known that if there is a random variable Q distributed such that its logarithm is $N(\mu,\sigma^2)$, then Q follows a log–normal distribution with mean $\exp(\mu + \sigma^2/2)$ and variance $[\exp(2\mu + \sigma^2/2) - \exp(2\mu + \sigma^2)]$. Here, if we assume the usual regression assumptions are satisfied, then $ln(B)$ should be approximately normally distributed and B should be approximately log–normally distributed. Hence for δ_{comb} the predictions were obtained from:

$$\hat{B}_i = \exp(x_i'\hat{\beta} + s_{reg}^2/2), \tag{9}$$

where: $s_{reg}^2 = MSE(1 + x_i'(X'X)^{-1}x_i)$, $x_i' = [1, \bar{A}_i^{*-1}, ln(\bar{N}_i^*), ln(\bar{H}_i^*)]$, $\hat{\beta}$ is the usual least squares estimate for the fitted model, MSE is the mean square error from the fitted model, and $X'X$ is the design matrix used to fit the model (i.e., $X'X$ comes from the 50 randomly selected plots). Here $\bar{A}_i^*$, $\bar{N}_i^*$, and $\bar{H}_i^*$ are the stand means over all plots, not the sample means of the five randomly drawn plots. This was done to minimize the dependence of $\hat{B}_i$ and $\bar{B}_i$. The bias correction in (9) was not used for δ_{p-2}^+. One advantage of this estimator is that the variance of the prediction is not needed, and hence the $X'X$ matrix is not required. Thus for δ_{p-2}^+ we simply used $\exp(x_i'\hat{\beta})$ for the model prediction. For δ_{comb}, τ^2 was estimated by $s_y^2 = \exp[2(x_i'\hat{\beta}) + 2\, s_{reg}^2] - \exp[2(x_i'\hat{\beta}) + s_{reg}^2]$, and the variance of $\bar{B}_i$ (σ^2) was the usual squared standard error for stand i. For δ_{p-2}^+, σ^2 was estimated by pooling the squared standard errors across stands. In other words, the variance of the $\bar{B}_i$'s were assumed to be equal in δ_{p-2}^+.[1] Bartlett's test for homogeneity of variances indicated that the per–stand variances of the B_{ij} were unequal at the α=0.05 level for the data from Companies 1 and 3, while the hypothesis of equality could not be rejected at the 0.05 level for data from Company 2.

The simulation was performed 1000 times. On each trial the model was fitted and 5 observations were then randomly drawn from each stand. Although the same fitted model was used for each company, the estimation procedures were not combined, i.e., estimates were calculated separately for each company. Thus the dimension of the problem (p) was 47, 9, and 10 for Companies 1, 2, and 3, respectively. The mean difference (**D**) and mean squared difference (**D²**) between the three estimators ($\bar{B}_i$, δ_{comb}, and δ_{p-2}^+) and $\bar{B}_i^*$, where the means are averaged over stands and the 1000 trials, are shown, by Company, in Table 1. Also shown in Table 1 are the standard errors over the 1000 trials of the two statistics (**D** and **D²**), the average maximum and minimum differences per stand (averaged over the 1000 trials), and the ratios of the **D** and **D²** values obtained for $\bar{B}_i$ and those for δ_{comb} and δ_{p-2}^+. As would be expected, the unbiased sample mean $\bar{B}_i$ performed best under the mean difference criterion. However, if one is willing to relax or disregard the unbiasedness requirement, the results demonstrate that one may realize practically significant, as well as statistically significant, savings (up to ~56% in squared error).

---

[1] An unequal variance version of δ_{p-2}^+ is currently under development.

Both δ_{comb} and $\delta_{\mathrm{p-2}}^{+}$ improve upon $\bar{\mathrm{B}}_{\mathrm{i}}$ under $\mathbf{D^2}$ for all three companies. The values indicate that bias of the model did not appear to be a serious problem for the data from Company 2, and thus δ_{comb} performed slightly better than $\delta_{\mathrm{p-2}}^{+}$. However, for Companies 1 and 3, the results indicate that bias of the model caused the performance of δ_{comb} to deteriorate. Although δ_{comb} was better than $\bar{\mathrm{B}}_{\mathrm{i}}$ for these companies, it was not better than $\delta_{\mathrm{p-2}}^{+}$, indicating that there were savings from accounting for the model bias for these two Companies. Note that in a *t*— or *z*—test for statistical significance, $\delta_{\mathrm{p-2}}^{+}$ would be found significantly lower than δ_{comb} under $\mathbf{D^2}$ for companies 1 and 3 at any conventional level of significance, while the differences between the two estimators for company 2 would not be found significant.

In terms of individual component (stand) coverage, the values in Table 1 show that the average maximum stand differences from δ_{comb} and $\delta_{\mathrm{p-2}}^{+}$ were always less than those for $\mathbf{X}$ for both statistics. Additionally, under $\mathbf{D^2}$, $\delta_{\mathrm{p-2}}^{+}$ was better than δ_{comb} for companies 1 and 3, while δ_{comb} was slightly better for company 2. These results indicate that $\delta_{\mathrm{p-2}}^{+}$ performed well on a stand—by—stand basis, as well as on an overall basis.

Recall that the variances of plot basal area were not homogeneous across stands for companies 1 and 3. Thus the superior performance of $\delta_{\mathrm{p-2}}^{+}$ for these two companies indicates that $\delta_{\mathrm{p-2}}^{+}$ is somewhat robust to the homogeneous variance assumption.

CONCLUSIONS

In this paper we have presented an alternative to the usual precision—weighted composite estimator. We have demonstrated, both theoretically and empirically, that the new estimator ($\delta_{\mathrm{p-2}}^{+}$) behaves nearly as well as the usual estimator (δ_{comb}) when the auxiliary information is unbiased, yet is superior to δ_{comb} when the auxiliary information is badly biased. In the latter case, the risk of δ_{comb} is unbounded from above, while the risk of $\delta_{\mathrm{p-2}}^{+}$ is bounded from above by the risk of the sample mean. Furthermore, as is evident in Figures 1a—f, and in the simulation results, the point at which the risk of δ_{comb} becomes larger than the risk of $\delta_{\mathrm{p-2}}^{+}$ is below the risk of the sample mean. Thus, even if the bias in the auxiliary information is not great enough to cause δ_{comb} to have higher risk than the sample mean, it is still possible for $\delta_{\mathrm{p-2}}^{+}$ to have lower risk than δ_{comb}. We conjecture that in many instances, investigators may possess auxiliary information which they are hesitant to use because they suspect that it may be biased. With $\delta_{\mathrm{p-2}}^{+}$, this is no longer a valid concern.

In order to use δ_{comb}, one must be able to estimate the variance of the auxiliary information. This implies that when the auxiliary information is represented by a regression model, one must have access to the $\mathbf{X'X}$ matrix used to calibrate the model (or, equivalently, $\mathbf{X'X^{-1}}$). This can often be a problem because authors do not usually publish $\mathbf{X'X}$. With $\delta_{\mathrm{p-2}}^{+}$, one does not need to specify the variance of the auxiliary information, and so not having access to the $\mathbf{X'X}$ matrix is not a problem.

Some words of caution are in order. First, $\delta_{\mathrm{p-2}}^{+}$ requires that one be interested in 3 or more problems simultaneously. The risk of $\delta_{\mathrm{p-2}}^{+}$ improves relative to the risk of δ_{comb} as the number of problems increases. Second, in $\delta_{\mathrm{p-2}}^{+}$ the variances of the sample means are assumed to be equal. Although this was not a problem in our simulation experiment, it may present difficulties in some instances and should be closely examined. If this assumption seems untenable in practice, a carefully designed simulation should be carried out to investigate the behavior of $\delta_{\mathrm{p-2}}^{+}$. Development of an unequal variance version of $\delta_{\mathrm{p-2}}^{+}$ is an important direction for future work.

Table 1. Mean difference and mean squared difference between assumed true basal area and estimated basal area, standard errors of mean over the 1000 simulations, average minimum and maximum stand differences, and ratio of each mean to the mean for X, by company and estimator.

	Company 1			Company 2			Company 3		
	X	δ_{comb}	δ_{p-2}^{+}	X	δ_{comb}	δ_{p-2}^{+}	X	δ_{comb}	δ_{p-2}^{+}
Mean Difference.	0.10	2.10	3.24	–0.01	–4.24	–4.20	0.07	–1.54	–2.62
std. error	0.06	0.05	0.06	0.15	0.14	0.17	0.15	0.12	0.09
avg. stand min.	–29.59	–23.36	–18.74	–21.90	–23.27	–23.08	–22.72	–20.57	–17.00
avg. stand max.	31.93	31.06	31.37	22.04	15.51	14.60	23.57	17.65	11.95
ratio	1.00	21.00	32.00	1.00	424.00	420.00	1.00	–22.00	–37.43
Mean Sq. Diff.	156.8	128.8	120.1	217.1	187.1	188.9	223.0	153.5	99.0
std. error	1.2	1.0	1.0	3.3	2.7	2.8	3.5	2.4	1.9
avg. stand min.	0.1	0.1	0.1	6.3	4.6	5.6	5.0	2.9	2.3
avg. stand max.	1325.7	1108.3	1076.8	820.5	694.5	704.5	898.1	623.7	378.4
ratio	1.00	0.82	0.77	1.00	0.86	0.87	1.00	0.69	0.44

REFERENCES

Green, E.J. and W.E. Strawderman. 1987. Cooperative for Alternative Parameter Estimation at Rutgers: Report I. Forestry & Wildlife Section, Cook College — Rutgers University, New Brunswick, NJ.

Green, E.J. and W.E. Strawderman. 1988. Combining inventory data with model predictions. p. 676—682. *In*: Forest Growth Modelling and Prediction, vol. 2. Proceedings, International Union of Forest Research Organizations Conference, Minneapolis, MN. Aug 23—27, 1987. USDA Forest Service General Technical Report NC—120. 1149 p.

Green, E.J. and W.E. Strawderman. 1989. A James—Stein type estimator for combining unbiased and possibly biased estimators. Manuscript submitted to: Journal of the American Statistical Association.

James, W. and C. Stein. 1961. Estimation with quadratic loss. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability 1:361—380. University of California Press, Berkeley, CA.

Lappi, J. and R.L. Bailey. 1988. A height prediction model with random stand and tree parameters: An alternative to traditional site index methods. Forest Science 34:907—927.

Lehmann, E.L. 1983. Theory of Point Estimation. Wiley and Sons, Inc., NY. 506 p.

Parresol, B.R., K.R. Dobelbower, and T.R. Dell. 1987. Honduran Pine Yield System. Institute for Quantitative Studies, USDA Forest Service, Southern Forest Experiment Station, New Orleans, LA.

The Variances of Some Compatible Estimators for The Components of Growth from Remeasured Point Samples

Francis A. Roesch Jr.
Mathematical Statistician
Institute for Quantitative Studies
USDA Forest Service, Southern Forest Experiment Station
701 Loyola Avenue, New Orleans, LA 70113

ABSTRACT

Since the inception of remeasured horizontal point samples a number of compatible estimators for the components of growth in the equation,

$$V_2 - V_1 = S + I - M - C,$$

have been proposed. This paper discusses the variances of a few of the estimators of S and I with respect to the change in the volume-to-basal area ratio, the relative length of the growth interval, the spatial distribution of the trees and the variance in growth of the trees within a size class.

Key Words. Point sampling, forest growth, forest inventory, compatible estimators, variance.

INTRODUCTION

The components of growth equation

$$V_2 - V_1 = S + I - M - C,$$

where :

V_i = the volume at time i,
S = survivor growth,
I = ingrowth,
M = mortality, and
C = cut

was presented by Meyer (1953). We can see that V_2 can be partitioned into 3 mutually exclusive components, S, I, and (V_1 - M - C). In Roesch et al. (1989) we presented a new compatible grouping of estimators for use with remeasured horizontal point samples which, if unmeasured time 1 parameters could be accurately estimated, used only the sample from each component for the estimators of each component. In Roesch (1988) this compatible grouping of estimators was shown to avoid the cost of compatibility noted by Flewelling (1981) (that of increased variance in the growth and/or yield estimators), at least in the case investigated. Therefore, the next question one might ask is " Is there a set of forest conditions that dictates which compatible grouping of estimators is the best to use?". If we limit our investigation to the possibly unbiased estimators and derive the basic expressions for the variances of the estimators, then this set of conditions, if it exists, will become obvious by what is left in the variance terms after reduction.

I will use the definitions of Martin (1982) for the six categories of trees encountered in remeasured point samples. By these definitions actual survivor trees can be sampled as either "s" trees or "n" trees. The former are measured at both points in time while the latter are measured only at time

2. Actual ingrowth trees can be sampled as either "i" trees or "o" trees with "i" trees being measured both times and "o" trees measured only at time 2. Mortality and cut trees will be represented by "m" and "c" respectively. In the same paper Martin presented notation utilizing these definitions which was augmented in Roesch et al. (1989). Simply, a prime indicates that the basal area divisor in the volume-to-basal area ratio (VBAR) is from the time period other than that subscripted. I will examine the variances of the survivor growth and ingrowth estimators in the three compatible groupings:

$$\hat{V}_2 - \hat{V}_1 = \hat{S} + \hat{I}_A - \hat{M} - \hat{C}$$

$$\hat{V}_2 - \hat{V}_1 = \tilde{S} + \hat{I}_p - \hat{M} - \hat{C}$$

$$\hat{V}_2 - V_1^{**} = S^{**} + \hat{I}_p - \hat{M} - \hat{C}.$$

Van Deusen et al. (1986) showed that these groupings differed by the placement of a common variate (ϵ), and we showed in Roesch et al. (1989) that the third grouping splits ϵ up into 2 parts, and applies one part to V_1^{**} and the other to S^{**}. Table 1 describes the estimators of S and I. In this paper I will ignore the estimators of instantaneous volume, mortality and cut.

We can define the indicator variable:

$$Y_{ij} = \begin{cases} 1 & \text{if tree } i \text{ is sampled from point } j \text{ as an "s" tree,} \\ 0 & \text{otherwise;} \end{cases}$$

Upper case M and N will represent the set of survivor trees and the set of ingrowth trees, respectively. In addition, I will define:

F = the basal area factor, the amount of basal area in square feet per acre represented by each "in" tree;

A = the area of the forest in acres;

b_{ji} = the basal area of tree *i* at time *j* in square feet, and;

V_{ji} = the volume of tree *i* at time *j* in square feet.

Table 1: The estimators of survivor growth and ingrowth calculated from remeasured horizontal point samples.

The estimators for survivor growth:

$\hat{S} = s_2' - s_1$	The traditional estimator (Grosenbaugh, 1958) as presented by Martin (1982).
$\tilde{S} = s_2 - s_1 + n_2$	The Van Deusen et al (1986) estimator.
$S^{**} = s_2 - s_1' + n_2 - n_1'$	An unbiased estimator which subtracts the estimated time 1 volumes of the "n" trees, presented in Roesch et al. (1989).

The estimators for ingrowth:

$\hat{I}_A = i + o + n_2 + s_2 - s_2'$	Martin's compatible estimator.
$\hat{I}_P = i + o$	The revised Purdue estimator.

Palley and Horwitz (1961) show the method of deriving the variance of the instantaneous point sampling estimators using basic geometric probability and Van Deusen et al. (1986) gave the probabilities of inclusion in the various samples on remeasured points (i, o, s, and n). They also gave the variance for $\tilde{S}$ relative to $\hat{S}$. The results I present here are more basic in that they do not involve other estimators.

Referring to Figure 1, we can define some other terms useful in the derivation of the variances of the estimators. The "a" annulus for a tree is the plot associated with basal area growth between times 1 and 2 and is equal in area to $(b_{2i} - b_{1i}) / F$. The "c" circle for a tree is the plot associated with the tree's basal area at time 1 and is equal in area to b_{1i} / F. The following areas of overlap determine the joint probability of inclusion for any two trees in the forest by a random point sample:

OAA_{il} = the area of overlap of the "a" annuli of trees *i* and *l*;

OCC_{il} = the area of overlap of the "c" circle of trees *i* and *l*;

OCA_{il} = the area of overlap of the "c" circle of tree *i* and the "a" annulus of tree *l*;

$OT2_{il}$ = the area of overlap of the time 2 plots of trees *i* and *l*.

$OT2C_{il}$ = the area of overlap of the time 2 plot of tree *i* and the "c" circle of tree *l*.

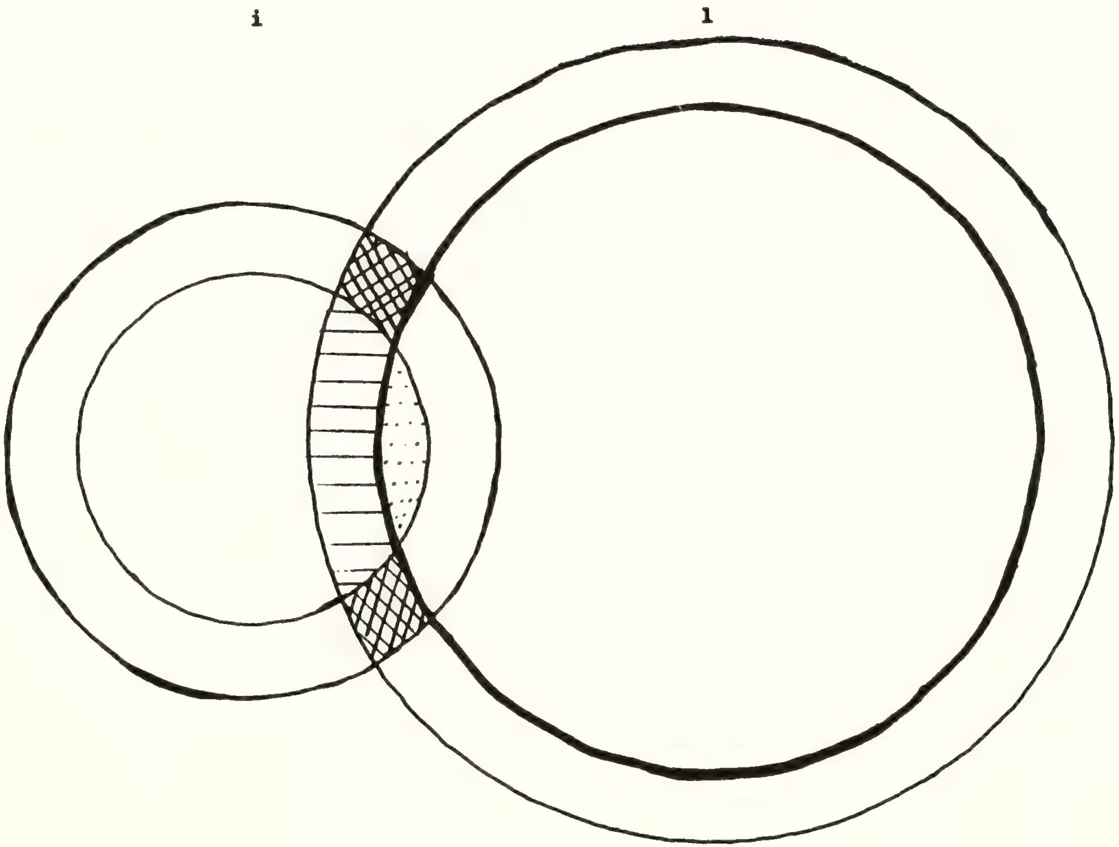


Figure 1: The areas of overlap of the remeasured variable-radius plots of trees i and 1. If the point falls in area OCC_{il} both trees are measured at time 1. If the point falls in areas OAA_{il} then both trees are measured only at time 2. When the point falls in area OCA_{il} tree i is measured at time 1 and tree 1 is only measured at time 2.

These overlap areas facilitate the calculation of the covariances of the indicator variables because the covariance of 2 Bernoulli random variables is simply the probability that they both equal 1 minus the product of their expected values. For example,

$$\mathrm{COV}[Y_{ij}, Y_{lj}] = E[Y_{ij} \cdot Y_{lj}] - E[Y_{ij}]E[Y_{lj}]$$

$$= \left[\frac{\mathrm{OCC}_{il}}{A} - \frac{b_{1i}}{FA} \cdot \frac{b_{1l}}{FA} \right]$$

So the variance of $\hat{S}$ is:

$$\mathrm{var}(\hat{S}) = \mathrm{var}(s_2' - s_1)$$

$$= \mathrm{var} \left[\frac{FA}{m} \sum_{j=1}^{m} \sum_{i \epsilon M} \left[Y_{ij} \left[\frac{V_{2i} - V_{1i}}{b_{1i}} \right] \right] \right] \tag{1}$$

Since the points are randomly located, we see that (1)

$$= \left[\frac{FA}{m} \right]^2 \sum_{j=1}^{m} \mathrm{var} \left[\sum_{i \epsilon M} \left[Y_{ij} \left[\frac{V_{2i} - V_{1i}}{b_{1i}} \right] \right] \right] \tag{2}$$

This allows us to use the covariance result for a linear combination of random variables and we see that (2)

$$= \left[\frac{FA}{m} \right]^2 \sum_{j=1}^{m} \left[\sum_{i \epsilon M} \sum_{l \epsilon M} \left[\mathrm{cov}(Y_{ij}, Y_{lj}) \left[\frac{V_{2i} - V_{1i}}{b_{1i}} \right] \left[\frac{V_{2l} - V_{1l}}{b_{1l}} \right] \right] \right]$$

$$= \left[\frac{1}{m} \right]^2 \sum_{j=1}^{m} \sum_{i \epsilon M} \sum_{l \epsilon M} [(V_{2i}V_{2l} - 2V_{2i}V_{1l} + V_{1i}V_{1l})((\mathrm{OCC}_{il}F^2A \,/\, b_{1i}b_{1l}) - 1)]$$

Similarly, we can derive the variances for all of the estimators of interest. I am ignoring the error associated with the necessity to separate the "o" trees and "n" trees in the estimators that require this separation, although this error may in some cases be significant. Detailed derivations for the estimators are available from the author upon request.

The composite estimator presented in Van Deusen et al. (1986) has the following variance:

$$\text{var}(\widetilde{S}) = \text{var}(s_2 - s_1 + n_2)$$

$$= \left[\frac{1}{m}\right]^2 \sum_{j=1}^{m} \sum_{i\epsilon M} \sum_{l\epsilon M} \Big[(V_{2i}V_{2l} / b_{2i}b_{2l})((\text{OCC}_{il}+\text{OAA}_{il}+2\text{OCA}_{il})F^2A - b_{2i}b_{2l})$$

$$+ 2(V_{1i}V_{2l} / b_{1i}b_{2l})(b_{1i}b_{2l} - (\text{OCC}_{il}+\text{OCA}_{il})F^2A)$$

$$+ (V_{1i}V_{1l} / b_{1i}b_{1l})((\text{OCC}_{il})F^2A - b_{1i}b_{1l})\Big]$$

The variance of S^{**} is slightly more complicated than that of $\widetilde{S}$ due to the predicted time 1 volumes of the "n" trees:

$$\text{var}(S^{**}) = \text{var}(s_2 - s_1' + n_2 - n_1')$$

$$= \left[\frac{1}{m}\right]^2 \sum_{j=1}^{m} \sum_{i\epsilon M} \sum_{l\epsilon M} \frac{1}{b_{2i}b_{2l}} \Big[V_{2i}V_{2l}((\text{OCC}_{il}+\text{OAA}_{il}+2\text{OCA}_{il})F^2A - b_{2i}b_{2l})$$

$$+ 2V_{2i}V_{1l}(b_{2i}b_{2l} - (\text{OCC}_{il}+\text{OCA}_{il})F^2A)$$

$$+ V_{1i}V_{1l}((\text{OCC}_{il})F^2A - b_{1i}b_{1l})$$

$$+ 2V_{2i}\hat{V}_{1l}((b_{2i}b_{2l} - b_{2i}b_{1l}) - (\text{OAA}_{il}+\text{OCA}_{il})F^2A)$$

$$+ \hat{V}_{1i}\hat{V}_{1l}((\text{OAA}_{il})F^2A - (b_{2i}-b_{1i})(b_{2l} - b_{1l}))$$

$$+ 2V_{1i}\hat{V}_{1l}((\text{OCA}_{il})F^2A - (b_{1i})(b_{2l} - b_{1l}))\Big]$$

The same techniques can be applied to determine the variances of the estimators of ingrowth. The variance of the revised Purdue estimator:

$$\text{var}(\hat{I}_P) = \text{var}(i + o)$$

$$= \left[\frac{1}{m}\right]^2 \sum_{j=1}^{m} \sum_{i\epsilon N} \sum_{l\epsilon N} [(V_{2i}V_{2l})((\text{OT2}_{il}F^2A / b_{2i}b_{2l}) - 1)]$$

The variance of Martin's estimator of ingrowth:

$$\mathrm{var}(\hat{I}_A) = \mathrm{var}(\, i + o + n_2 + s_2 - s_2')$$

$$= \left[\frac{1}{m}\right]^2 \sum_{j=1}^{m} [\sum_{i \epsilon M,N} \sum_{l \epsilon M,N} (V_{2i}V_{2l} / b_{2i}b_{2l}) ((OT2_{il})F^2A - b_{2i}b_{2l})$$

$$+ \sum_{k \epsilon M} \sum_{h \epsilon M} (V_{2k}V_{2g} / b_{1k}b_{1g}) ((OCC_{kg})F^2A - b_{1k}b_{1g})$$

$$- \; 2 \sum_{h \epsilon M,N} \sum_{f \epsilon M} (V_{2h}V_{2f} / b_{2h}b_{1f}) ((OT2C_{hf})F^2A - b_{2h}b_{1f})]$$

CONCLUSIONS

Now that the variances of these estimators are presented in a uniform terminology, we can begin to examine the optimality of the estimators based on the conditions present in any one forest. We can see that the conditions affecting the relative variances of these estimators are:

(1) the volume to basal area ratios at each point in time, and the change in that ratio;

(2) the length of the measurement interval relative to the age of the trees at the second inventory;

(3) the spatial distribution of the respective size classes at time 2; and,

(4) the variance of growth since time 1 within each size class at time 2, and the various interrelationships of these conditions.

We can see (1) is true by the VBAR terms appearing in the variance terms. (2) becomes obvious by the b_{1i} / b_{2i} terms appearing in the formulas, while the overlap terms support both (2) and (3). Although condition (4) is evident in the $\hat{V}_{1i}$ terms of S^{**}, this variance also affects the distinction between ongrowth and nongrowth. That the interrelationships are important are evidenced by our inability to algebraically separate the 4 conditions. One can see that even the introduction of set theory to eliminate the overlap terms will not aid in the separation of these four conditions, although assumed (or known) spatial distributions would help. We are currently conducting simulations intended to investigate the relative effect of these factors on the various component estimators. The results of these simulations should be available soon.

REFERENCES

Flewelling, J.W. 1981. Compatible estimates of basal area and basal area growth from remeasured point samples. Forest Sci. 27(1):191-203.

Grosenbaugh, L.R. 1958. Point-sampling and line-sampling: Probability theory, geometric implications, synthesis. USDA Forest Service Occasional Paper 160. 34 p.

Martin, G.L. 1982. A method of estimating ingrowth on permanent horizontal sample points. Forest Sci. 28(1):110-114.

Meyer, H.A. 1953. Forest Mensuration. Penn Valley Publishers, Inc. State College, Pa. 357 p.

Palley, M.N., and L.G. Horwitz. 1961. Properties of some random and systematic point sampling estimators. Forest Sci 7(1): 52-65.

Roesch, F.A. Jr. 1988. Compatible estimators for the components of growth and instantaneous volume calculated from remeasured horizontal point samples. Unpublished PhD Dissertation, Rutgers University, New Brunswick, N.J.

Roesch, F.A. Jr., E.J. Green and C.T. Scott. 1989. New compatible estimators for survivor growth and ingrowth from remeasured horizontal point samples. Forest Sci. 35(2):281-293.

Van Deusen, P.C., T.R. Dell and C.E. Thomas. 1986. Volume growth estimation from permanent horizontal points. Forest Sci. 32(2):415-422.

THE TRUNCATED LAW

R.B. Chevrou
Statistician
National Forest Inventory Service
34970 LATTES - FRANCE

ABSTRACT

The truncated law is a diameter distribution that better fits real data than the De Liocourt's law for irregular stands as well as for regular stands. With parameters α, β and the proportional coefficient M, the number of trees N(*,i) in the diameter class i, is :

$$N(*,i) = \frac{1}{(1+\alpha)} \left\{ N(*,i-1) - \alpha M e^{-(1+\alpha)\beta} \frac{\beta^i}{i!} \right\}$$

In balanced regular stands, β can be related to maximum age. When β tends to infinity, $1+\alpha$ is the usual q-ratio.

The Truncated Law Formula

The truncated law is an empirical diameter distribution. Let's u be the diameter class width, and d = u i the diameter of the class i.

N(*,i) is the number of trees in the diameter class i, M a proportional coefficient, α and β the two parameters of the truncated law :

$$(1) \quad N(*,i) = \frac{1}{(1+\alpha)} \left\{ N(*,i-1) - \alpha M e^{-(1+\alpha)\beta} \frac{\beta^i}{i!} \right\}$$

from CHEVROU (1989a) and (1989b).

The formula (1) can be established from the following distribution, where x is an anonymous and continuous variable taking values in the interval $(0,\beta)$:

$$(2) \quad N(x,i) = \alpha M e^{-(1+\alpha)x} \frac{x^i}{i!}$$

Integrating (2) by parts, it comes :

$$(3) \quad N(*,i) = \int_0^\beta N(x,i) \, dx = \frac{1}{1+\alpha} \left\{ N(*,i-1) - N(\beta,i) \right\}$$

with (1) as the normal form of the truncated law.

From (1) and (2), it comes, see CHEVROU (1989b) :

For Mt, the total number of standing trees :

$$(4) \qquad Mt = M \left\{ 1 - e^{-\alpha\beta} \right\}$$

For dm, the mean diameter of these Mt trees :

$$(5) \qquad dm = \left\{ \frac{u}{\alpha} \right\} \left\{ 1 - \alpha\beta \frac{e^{-\alpha\beta}}{1 - e^{-\alpha\beta}} \right\}$$

Taking ß infinite leads to the usual De Liocourt's law, with the De Liocourt's parameter q, or q-ratio, equal to $q = 1+\alpha$.

Usually we do not tally smaller trees, so that i has a minimum value $k > 0$. The above values can be easily computed with $i \geq k$. Formulas are cumbersome but for ß infinite, that is for the De Liocourt's law, to get from Chevrou (1989b) :

For M(k), the number of trees of size $d \geq k\,u$:

$$(6) \qquad M(k) = \frac{q\, N(*,k)}{(q-1)} \qquad \text{and : } N(*,k) = \frac{(q-1)\, M(k)}{q}$$

For dm(k), the mean diameter of these M(k) trees :

$$(7) \qquad dm(k) = k\,u + \frac{u}{q-1} \qquad \text{and : } \quad q = 1 + \frac{1}{[\, dm(k) \,/\, u \,] - k}$$

Fitting The Truncated Law

The truncated law can be fitted to real data by the minimum sum of squares technique.

This fitting can be based on real values of numbers of trees per diameter classes, or on real values of volumes per diameter classes, regarding the mean tree volume in a class as a convenient weigth.

Working with numbers of trees leads to a closer fit for small diameters where numbers are larger, while working on volumes leads to a closer fit for medium size classes where volumes are larger. Such a fit on volume obviates the problem related to undesirable regeneration in medium age stands with many young small trees in old stands which bring some trouble in fitting the truncated law as well as De Liocourt's.

An example is given in table 1 from CHEVROU (1989a).

Table 1. Fir (Abies Alba) stands in balanced public forests, Vosges, Northeast of France.

Diameter classes cm	Real data		Best fitted q-ratio		Truncated law	
	Numbers	Volumes m^3	Numbers	Volumes m^3	Numbers	Volumes m^3
20	406 048	107 011	489 102	128 899	402 434	106 059
25	328 262	147 934	353 455	159 287	332 733	149 949
30	273 769	215 602	255 429	201 159	270 661	213 154
35	216 861	257 553	184 589	219 225	214 947	255 279
40	164 659	285 308	133 395	231 136	165 395	286 584
45	118 886	270 333	96 400	219 202	122 509	278 572
50	81 769	235 196	69 664	200 378	86 916	249 999
55	66 545	230 884	50 344	174 673	58 861	204 222
60	44 205	184 136	36 381	151 545	37 973	158 177
65	15 841	76 840	26 292	127 535	23 316	113 099
70	12 647	73 537	19 000	110 477	13 625	79 222
75	5 540	35 962	13 730	89 126	7 581	49 209
80	7 225	50 871	9 922	69 860	4 020	28 302
85	2 449	20 901	7 171	61 201	2 033	17 354
90	716	7 753	5 182	56 112	983	10 641
TOTAL	1 745 422	2 199 821	1 750 056	2 199 815	1 743 987	2 199 822
R^2	--	--	0.95	0.82	0.999	0.97

Source : Results from the second national forest inventory in 1981.

Relating The Variable x To Age a

The anonymous variable x can be related to age a, with $x = a / c$, c being a convenient value seen as an age class width. Chevrou (1989a).

It permits to bring some meaning to the truncated law parameters, but the distribution (2) may not fit exactly real age data.

To the maximum value ß of x, corresponds a maximum value MA of the age, with MA = ß c. This maximum age may be an exploitable age.

The distribution becomes :

$$N(a,i) = \frac{\alpha M}{c} e^{-(1+\alpha)a/c} \frac{(a/c)^i}{i!} \qquad (8)$$

The number of trees of age a is :

$$N(a,*) = \sum_{i=0}^{\infty} N(a,i) = \frac{\alpha M}{c} e^{-\alpha a/c} \qquad (9)$$

This formula may correspond to balanced regular stands if $\alpha \geq 0$.

From (8) it comes, see Chevrou (1989a) and (1989b) :

For d(a), the mean diameter of trees of same age a :

$$d(a) = \frac{u a}{c} \qquad (10)$$

The mean diameter DMA of trees having the maximum age MA is :

$$DMA = d(MA) = \frac{u\,MA}{c} = \beta u$$

So, to the exploitable age MA it corresponds a clear exploitable diameter DMA, which is the mean diameter of trees located in the stands of exploitable age MA.

It can be seen that MA and DMA come as familiar parameters in the truncated law, while the value given to maximum diameter MD comes as an arbitrary parameter in the usual De Liocourt's law, MD being a value to truncate the law more or less arbitrarily to the right.

For a(d), the mean age of trees having the same diameter d = u i:

$$(11) \quad a(d) = \frac{c\ (i+1)\ N(*,i+1)}{N(*,i)} \le \frac{c\ (i+1)}{(1+\alpha)}$$

a(d) has a maximum value MA = ß c, so that N(*,i+1)/N(*,i) tends to ß/(i+1) when d tends to infinity.

There is some symetry between age classes and diameter classes seen as variables, and corresponding results. Inverting the parts played by relative age a/c and relative diameter d/u, with diameter limited to a maximum value MD and unlimited age, would give similar results which do not work so well.

Annual Crop

N(MA,i), or N(ß,i) in (3), can be seen as the final crop of trees of diameter d = u i and of exploitable age MA.

Thinning at age a cut T(a,*) trees with :

$$T(a,*) = N(a,*) - N(a+1,*) \approx (\alpha/c)\ N(a,*)$$

When the thinning type ratio is 1, as it is in irregular stands, the number of thinned trees of age a and with diameter d = u i, is :

$$T(a,i) \approx (\alpha/c)\ N(a,i)$$

and the annual crop of trees of diameter d = u i can be computed adding N(MA,i) and T(a,i) to get :

$$(12) \quad C(*,i) = \left\{ \frac{N(*,i-1) - N(*,i)}{c} \right\} \qquad \text{from CHEVROU (1989b)}$$

From (8), the basal area of trees of age a is G(a) :

$$(13) \quad G(a) = \frac{\pi\ u^2}{4}\ \frac{\alpha M}{c}\ e^{-\alpha\ a/c}\ (a/c)\ (1 + a/c)$$

From the value of G(a) and the value of thinnings between age 0 and age a, the basal area yield at age a is given by the following formula of a familiar type :

$$(14) \quad F(a) = K \left\{ 1 - \left[1 + \frac{\alpha a/c}{1+\alpha/2} \right] e^{-\alpha a/c} \right\}$$

The Meanings Of Parameters α, ß, u And c

The parameter α is related to the thinning ratio (or natural mortality). The parameter ß is related to maximum value of age or the value of exploitable age, and to mean diameter DMA of oldest trees having maximum age MA.

The parameter u is a diameter class width, and the parameter c can be seen as an age class width. Their ratio u/c is meaningful as expressing the mean value of diameter growth, that is as expressing a site index, c being related to site for given u. Strictly speaking the values of u and of c are related only on fitting real age data and they come as additional parameters in the relationship (8).

From (7), the value of the q-ratio depends only on the value of dm/u, or dm(k)/u, that is on the mean diameter of all standing trees ; it can be seen, from (5) and (8), that the q-ratio depends on site through the value of u/c, on sylviculture through the value of α, and on the exploitable age through the value of ß.

There is a big difference between regular stands and irregular ones in relation with thinning type ratio, or the ratio of mean diameter of thinned trees of given age a on mean diameter of standing trees of same age a. In irregular stands, where tree age is not known, this ratio is 1. In regular stands this ratio may be different from 1. When it is less than 1, thinnings cut smaller trees of age a ; for given site index, or given yield F(a), they increase apparent diameter growth in regular stands, and they decrease the values of MA and c for given value of ß and DMA.

This makes difficult to compare data for regular stands to those for irregular stands.

Ages In Irregular Stands

Let a stand where the truncated law (1) and the relationship (8) are valid.

Let θ the diameter annual growth of a given tree, and θm = u/c.

Let p(i) the probability of a tree with diameter d = u (i-1), to be cut during a period of time of length c.

From (12) it comes, approximatively :

$$(15)\qquad p(i) = \frac{C(*,i)}{N(*,i-1)} = 1 - \frac{N(*,i)}{N(*,i-1)}$$

The probability for the tree not to be cut during this period of time, is :

$$(16)\qquad q(i) = 1 - p(i) = \frac{N(*,i)}{N(*,i-1)}$$

A tree with diameter growth θ, stays in a diameter class during a period of time equal to c θm/θ, and the probability for this tree not to be cut while staying in the diameter class i-1, is q(θ,i) :

(17) $$q(\theta, i) = q(i)^{\theta m/\theta}$$

(17) leads to the probability Q(θ,i), or Q(a,i), for a tree to reach diameter d = u i, and age a = c i θm/θ :

(18) $$Q(\theta, i) = \left\{ \frac{N(*, i)}{N(*, 0)} \right\}^{\theta m/\theta} \quad \text{Or} \quad Q(a, i) = \left\{ \frac{N(*, i)}{N(*, 0)} \right\}^{a/ci}$$

It can be shown, from (1), that Q(a,i) ≤ Q(a,i-1). Obviously, from (15), trees with greater diameter growth reach earlier larger diameter classes where the probability to be cut is larger. The probability for a tree to reach old ages should be greater for trees having smaller diameter growth.

Nevertheless, trees cannot reach infinite ages. Sylvicultural practices, as well as nature, eliminate older trees for various reasons, specially in smaller diameter classes. Real data show that relatively older trees are common in middle diameter classes, and they are a few in smaller and larger diameter classes. See CHEVROU (1989a).

So the relationships (18) do not work well, while the truncated law fits irregular stands data.

Conclusion

There is a lot to do for understanding age and diameter structures in irregular stands.

Up to now, and in most cases, irregular stands have been studied through theories related to regular ones.

The truncated law better fits the data than De Liocourt's, and it may be an improvement, but not a full answer.

REFERENCES

Chevrou, R.B., 1989a, La loi tronquée de De Liocourt, proposed for publication to Ann. Sc. Forest. (May 1989).

Chevrou, R.B., 1989b, The De Liocourt's truncated law, proposed for publication to Forest Sci. (June 1989).

Sampling Mixed Stands Using Product Estimator

Krishna P. Rustagi
Associate Professor
Center for Quantitative Science
College of Forest Resources
University of Washington, HR-20
Seattle, WA 98195

ABSTRACT

In this sampling scheme, per acre volume by species and its standard error are estimated using the product rule for two random variables: the basal area per acre from the tree count data, and average of the volume-basal-area-ratio of a subsample of point sampled trees. Both the volume estimate and its sampling error are unbiased. In mixed stands, species can be aggregated and the sampling error of the combined volume estimate computed. For a given sampling effort, this method is more efficient than the traditional count and volume plot method, and is ideally suited for inventorying mixed stands.

INTRODUCTION

Bell, Iles and Marshall (1983) proposed estimation of stand volume as the product of average stand basal area $(\overline{ba})$, from tree count on variable plots, and the average of volume-basal-area-ratios $(\overline{vbar})$ of a subsample of all **in** trees included in the first sample. We shall call this estimator the 'product estimator' as the per acre volume is estimated as the product of two random variables (i.e., $\overline{ba}$ and $\overline{vbar}$). They assumed statistical independence between $\overline{ba}$ and $\overline{vbar}$ and used the product rule by Goodman (Des Raj 1968) to compute the variance of their volume estimator.

It will be shown that, for all practical purposes, $\overline{ba}$ and $\overline{vbar}$ are statistically independent, and the use of the product rule in computing the sampling variance is appropriate. An alternative method for subsampling of trees is suggested which provides for better control of the subsampling intensity by individual species and/or tree size. A procedure is outlined for aggregating volumes and their sampling variances across species and strata. The results of sampling using the product estimator are compared with Johnson's (1961) double sampling ratio estimator in two stands. Finally, an example of the use of the product estimator in inventorying a mixed stand is presented.

POINT SAMPLING AND THE PRODUCT ESTIMATOR

An estimate of the per acre volume ($\hat{V}$) in point sampling is given by:

$$\hat{V} = F * \frac{\left[\sum_i \sum_j vbar_{ij}\right]}{n} \tag{1}$$

where F is the basal area factor, $vbar_{ij}$ is the volume-basal-area-ratio of the jth **in** tree on the ith variable plot, and n is the total number of variable plots in the sample.

Equation (1) may be rewritten as:

$$\hat{V} = F * \left(\frac{\sum_i m_i}{n}\right) * \frac{\{\sum \sum vbar_{ij}\}}{\sum_i m_i}$$

$$= \overline{ba} * \overline{Vbar} \tag{2}$$

where m_i is count of **in** trees on *i*th plot; $\overline{ba}$ is the average basal area per acre obtained from the average tree tally per plot; and $\overline{Vbar}$ is the average of the *vbar*s of all **in** trees on all plots in the sample.

The true stand volume per acre $(\overline{V})$ can be expressed as a product of the true basal area per acre $(\overline{BA})$ and the ratio of $\overline{V}$ to $\overline{BA}(\overline{VBAR})$.

$$\begin{aligned} \overline{V} &= \overline{BA}*\left(\frac{\overline{V}}{\overline{BA}}\right) \\ &= \overline{BA}*\overline{VBAR} \qquad (3) \end{aligned}$$

The similarities between (2) and (3) may be noted. Palley and Horwitz (1961) have already shown that both $\hat{V}$ and $\overline{ba}$ are unbiased estimators of $\overline{V}$ and $\overline{BA}$ respectively. It follows, therefore, that $\overline{Vbar}$, (a simple arithmetic mean of the $vbar_{ij}$ in (2)), must also be an unbiased estimator of $\overline{VBAR}$.

STATISTICAL INDEPENDENCE BETWEEN $\overline{ba}$ AND $\overline{Vbar}$

Even though both $\overline{ba}$ and $\overline{Vbar}$ are derived from the same data, the manner in which they are computed, and the information they provide is quite different. Implicit in the statistical dependence between $\overline{ba}$ and $\overline{Vbar}$ is some form of association between the tree count per plot (m_i) and the mean of *vbars* of the trees on that plot $(\overline{vbar}_i)$. Therefore, if there is no association between m_i and $\overline{vbar}_i$, the statistical independence between $\overline{ba}$ and $\overline{Vbar}$ is assured. The scatterplot of m_i and $\overline{vbar}_i$ in the two even-aged stands examined for this purpose showed no trend.

Even if m_i and $\overline{vbar}_i$ were statistically dependent, which could happen if a stand consisted of clusters of young and mature trees, statistical dependence between $\overline{ba}$ and $\overline{Vbar}$ would be weak, provided a reasonably large sample of variable plots is taken to estimate basal area per acre. Therefore, $\overline{ba}$ and $\overline{Vbar}$ in (2) may be considered independent for all practical purposes.

As a matter of fact, the product estimator would be far more efficient than the traditional fixed or variable plot sampling, if there is positive association between m_i and $\overline{vbar}_i$. The following example illustrates this:

Suppose m_i and $\overline{vbar}_i$ are statistically independent in a stand. We could rearrange the $vbar_{ij}$ from this data set by reassigning trees with smaller *vbars* to plots with smaller tree count, and trees with larger *vbars* to plots with larger tree count. This rearrangement with strong positive association between m_i and $\overline{vbar}_i$ would not alter the $\overline{ba}$, $\overline{Vbar}$ or $\hat{V}$, and therefore, the traditional variable plot sampling or product estimator would give the same estimate for the stand volume per acre. There would be no change in the variance of $\overline{ba}$ and $\overline{Vbar}$, with the result the product estimate would be equally efficient in these two stands. However, as the estimated volume per acre $(v_i = F*m_i*\overline{vbar}_i)$ on the *i*th plot in the new stand would be more variable due to positive association between m_i and $\overline{vbar}_i$, product estimator would be more efficient than the traditional sampling with or without double sampling. It should be noted, however, that the product estimator would be less efficient if m_i and $\overline{vbar}_i$ are negatively associated.

If, instead of measuring $vbar_{ij}$ on all **in** trees, *vbars* are measured on a random subsample, the average of these *vbars* $(\overline{vbar})$ -- being an unbiased estimator of $\overline{Vbar}$ -- is also an unbiased estimator of $\overline{VBAR}$. This means that an unbiased estimate of stand volume may be obtained as the product (Bell, Iles and Marshall 1983):

$$\bar{v} = \overline{ba} * \overline{vbar} \quad \cdot \qquad (4)$$

If $\overline{ba}$ and $\overline{Vbar}$ are statistically independent, so also would be $\overline{ba}$ and $\overline{vbar}$. Therefore, a valid estimate of the variance $\bar{v}$ may be obtained by the product rule (Goodman in Des Raj 1968):

$$var(\bar{v}) = \overline{ba}^2 * var(\overline{vbar}) * var(\overline{ba}) * \overline{vbar}^2 - var(\overline{ba}) * var(\overline{vbar}) \quad \cdot$$

As the last term on the right is much smaller than the other two terms, it is generally ignored resulting in a somewhat simpler but conservative expression:

$$var(\bar{v}) = ba^2 * var(\overline{vbar}) + \overline{vbar}^2 * var(\overline{ba}) \quad \cdot \qquad (5)$$

If volume estimates and their variances are needed for individual species and/or size classes, $\overline{ba}$ and $\overline{vbar}$ as well as their variances will have to be computed for each category. These may then be aggregated as needed. The following procedures are suggested for computing variance of aggregated volumes:

a. For computing variance of the combined average basal area (for more than one species or size class), within same stand/stratum, the tree count for all species and/or sizes should be aggregated on each plot first and then the variance of the combined average basal area computed.

b. For computing variance of the aggregated mean vbar, within one stand/stratum, the rules of double sampling for stratification should be used. With relative frequencies of point sampled trees as the strata weights, the following simplification of the variance formula for double sampling for stratification by Rao (Cochran 1977) is suggested:

$$var(\overline{vbar}) = \frac{\sum_j \{n_j * (n_j - 1) * var(\overline{vbar}_j)\}}{\left\{\left(\sum_j n_j\right) * \left(\sum_j n_j - 1\right)\right\}} + \frac{\sum_j \{n_j * (\overline{vbar}_j - \overline{vbar})^2\}}{\sum_j n_j \left(\sum_j n_j - 1\right)}$$

where n_j is the total count of **in** trees of *j*th species and/or size class, $var(\overline{vbar}_j)$ is the variance of the subsample mean $\overline{vbar}_j$ for the *j*th species/size class, and $\overline{vbar}$ is the aggregated mean of the *vbar*s.

c. Once the variance of aggregated $\overline{ba}$ and $\overline{vbar}$ has been computed, the variance of combined volume may be computed using (5).

PROPOSED SAMPLING AND COMPUTATIONAL PROCEDURE

Bell, Iles and Marshall (1983) have suggested using a larger basal area factor for subsampling of trees for *vbar*. However, their approach for subsampling has two potential drawbacks: one, it may not be easy to find a device with the right BAF, and two, there will be no control on the subsample size by species in mixed stands. Both of these can be rectified by taking a systematic 1-in-k_j subsample of trees with random start, where k_j is species dependent. Besides simplifying field work, the main advantage of this method is in providing better control on the fraction of all **in** trees by species which are in the subsample.

The method proposed here differs slightly from that of Bell, Iles and Marshall (1983). The following steps are involved in point sampling with product estimator for multiple species stand:

a. *Determination of the needed sample size.* When working with single species, the number of variable plots to be taken, and the number of trees to be subsampled could be obtained using method of Bell et al. In the case of mixed stands, the subsampling fraction may be computed for the principal species, and as far as practicable, the number of trees to be subsampled for other species may be fixed after taking into account their relative importance and preponderance.

b. *Establishment of variable plots on a systematic grid with a random start.* The advantage of systematic grid is in post-stratification, if the strata boundaries are not known in advance.

c. *Count of all* **in** *trees per plot separately by species/size class.* A running total tally, by species/size class should also be maintained to identify trees in the subsample. Unless a stand table is also to be constructed, no diameter measurements need be taken.

d. *Careful measurement of all trees in the subsample for determination of their vbars.* This may involve measurements of breast height diameter, total tree height and some measure of stem form.

e. *Calculation of mean vbar and its variance for each species/size class.* This is computing individual *vbar*s from the field data and appropriate volume tables, and finding their means and sampling variances.

f. *Computation of aggregated vbar.* The aggregated mean *vbar* is computed using stratification. The relative frequencies of **in** trees by species/size classes provide the strata weights.

g. *Computation of the variance of the aggregated* $\overline{vbar}$. The variance of the aggregated mean *vbar* may be computed using (5), which is based on double sampling for stratification.

h. *Computation of individual and aggregated volumes and their variances.* Once the mean basal area and associated mean *vbar* have been computed, (4) and (5) may be used to obtain volume estimates and their variances.

TWO EXAMPLES

Two examples are presented. In the first example, two even-aged single species stands are used to show the superiority of the product estimator over Johnson's ratio estimator. In the second example, inventory results of one multispecies stand are summarized by species and strata.

Single Species Stands (Pack Forest). Two Douglas fir stands were sampled with 120 and 137 variable plots established in the two Douglas fir stands and all included trees were measured for *vbar*. The summarized data from these two stands is in Table 1.

Table 2 shows the relative performance of the product estimator vis-a-vis Johnson's ratio estimator. For Johnson's ratio estimator, every 10th plot was the volume plot resulting in 10 different samples of volume plots. For the product estimator every 20th tree was taken in the subsample resulting in 20 different subsamples. The sampling error of the product estimator is smaller and has narrower range in spite of the fact that only about half as many trees were measured for *vbar* for product estimator.

Multispecies stand (Bigbeef Creek Forest). This 244 acre forest was post-stratified into three strata: (1) Conifer, (2) Mixed, and (3) Hardwoods. Separate estimates were needed for Douglas fir, Western hemlock, Western red cedar and Red alder both for sawlogs (in boardfeet) and smallwood (in cubic feet). Total aggregated volume and its standard error were needed for the conifer sawtimber and its standard error. One variable plot per acre was established and the subsampling fraction was 1 in 10 for Douglas fir and Western red cedar, and a more flexible intensity for Western hemlock.

Tables 3 and 4 show the data summary for conifer sawlogs only. In Table 3, the volumes have been aggregated first for each strata and then combined for the whole forest. In Table 4, the volumes have been aggregated first by species. It may be noted that the computed variance for the total saw timber volume came out to be essentially same (755 *vs.* 764) when the data was aggregated two different ways.

DISCUSSION

There are two distinct advantages in point sampling with product estimator:

1. It enables more efficient use of the total sampling effort. The major source of variation in per acre volume is due to variation in basal area per acre. As can be seen in Table 1, the coefficient of variation in basal area and volume are almost equal in both stands. Therefore, it makes more sense to take a large sample of variable plots and a smaller subsample of trees. A large sample of variable plots ensures statistical independence between $\overline{ba}$ and $\overline{vbar}$.

2. There is operational flexibility in setting up the subsampling intensity for different species in mixed stands. For example, in the Bigbeef creek area, the subsampling intensity could have been different not only in different strata, but also for different species in the same stratum, resulting in increased precision of the sample estimate, without increase in the number of trees actually measured.

As relatively few trees need be measured for *vbar*, effort could be directed in obtaining more precise estimates of individual tree volumes. Rustagi and Loveless (1989) have suggested a simple measure of stem form inclusion of which in the volume/biomass prediction model is likely to result in substantial reduction in the residual error of prediction.

BIBLIOGRAPHY

Bell, J.F., Iles, K., and Marshall, D.D., 1983: Balancing the ratio of tree count only sample points and VBAR measurements in variable plot sampling. Proc. Inter. Conf., "Renewable resource inventories for monitoring changes and trends", (Bell & Atterbury eds.) OSU, Corvallis, OR. SAF 83-14:699-702.

Cochran, W.G., 1977: Sampling techniques, 3rd ed. John Wiley.

Des Raj, 1968: Sampling theory. McGraw-Hill series in probability and statistics.

Johnson, F. 1961: Standard error of volume per acre under point sampling. USDA Forest Service Research Note PNW RN-201.

Palley, M.N., and Horwitz, L.G., 1961: Properties of some random and systematic point sampling estimators. Forest Science 16(3):240-246.

Rustagi, K.P., and R. Loveless, 1989: Improved cubic volume prediction using a new measure of form-factor. Jour. of For. Ecol. and Mgmt. (In press)

Table 1. Summary of variable plot data collected from two Douglas fir stands.

Variable	Count (Trees/Plots)	Mean	Standard Deviation	Standard Error
		Stand #1		
BA (ft^2)	120 plots	157.5	52.0	4.7
VBAR (ft)	945 trees	39.5	9.7	.32
VOLUME (ft^3)	120 plots	6228	2274	207.5
Volume (ft^3)	120/945	6228		191.5
		Stand #2		
BA (ft^2)	137 plots	128.8	62.2	5.3
VBAR (ft)	882 trees	36.4	8.4	.28
VOLUME (ft^3)	137 plots	4690	2505	214.1
Volume (ft^3)	137/882	4690		195.8

(VOLUME refers to the individual plot volume data summary and Volume refers to the product estimate of the volume per acre.)

Table 2. Comparison of simulated systematic sampling in two Douglas fir stands using ratio type double sampling, and subsampling for product estimator.

Double Sampling Ratio Estimate (1 in 10 plots)				Product Estimates With 1 in 20 Subsamping of Trees			
Estimated Volume (ft^3/ac)		Standard Error (ft^3/ac)		Estimated Volume (ft^3/ac)		Standard Error (ft^3/ac)	
Mean	Range	Mean	Range	Mean	Range	Mean	Range
			Stand #1				
6229	6078-6471	314.8	206.8-383.0	6228	6035-6455	252.2	219.9-285.1
			Stand #2				
4681	4123-5020	270.1	216.1-365.8	4690	4404-4833	229.4	219.6-236.4

Table 3. Bigbeef Creek Forest sawtimber bdft volume summary. (Volume has been aggregated first by strata.)

Stratum I

		Ba/Acre				← VBAR →		Volume per Acre	
Spp	Area (Acre)	Mean	SE	Tree Tally	Trees	Mean	SE	Mean	SE
DF	94	102.55	7.54	482	48	217.3	6.6	22285	1773
WH	94	6.38	1.55	30	8	202.0	19.2	1289	336
RC	94	21.28	4.04	100	9	147.1	10.4	3130	634
Total	94	130.2	7.02	612	65	205.1	5.6	26704	1616

Stratum II

		Ba/Acre				← VBAR →		Volume per Acre	
Spp	Area (Acre)	Mean	SE	Tree Tally	Trees	Mean	SE	Mean	SE
DF	85	27.06	3.82	115	10	197.5	13.5	5344	838
WH	85	10.59	2.40	45	6	263.0	25.4	2785	686
RC	85	29.41	4.56	125	12	118.5	14.7	3485	692
Total	85	67.06	5.38	285	28	173.2	9.8	11614	1141

Stratum III

		Ba/Acre				← VBAR →		Volume per Acre	
Spp	Area (Acre)	Mean	SE	Tree Tally	Trees	Mean	SE	Mean	SE
DF	65	8.00	2.88	26	2	198.0	0.3	1584	570
WH	65	1.23	0.97	4	2	203.3	9.0	250	198
RC	65	10.15	2.20	33	4	110.8	17.9	1125	304
Total	65	19.38	3.62	63	8	152.7	10.9	2959	591

Total Sawtimber Volume

		Ba/Acre				← VBAR →		Volume per Acre	
Strat	Area (Acre)	Mean	SE	Tree Tally	Trees	Mean	SE	Mean	SE
I	94	130.21	7.02	612	65	205.1	5.6	26704	1616
II	85	67.06	5.38	285	28	173.2	9.8	11614	1141
III	65	19.38	3.62	63	8	152.7	10.9	2959	591
Total	244	78.7	3.43	960	101	192.2	4.7	15122	755

Table 4. Bigbeef Creek Forest sawtimber bdft volume summary. (Data has been aggregated first by species.)

(Douglas Fir)

						← VBAR →		Volume per Acre	
Strat	Area (Acre)	Basal Mean	Area/Ac SE	Tree Tally	Trees	Mean	SE	Mean	SE
I	94	102.56	7.54	482	48	217.3	6.6	22286	1773
II	85	27.06	3.82	115	10	197.5	13.5	5344	838
III	65	8.00	2.88	26	2	198.0	0.3	1584	570
Total	244	51.07	3.29	623	60	212.8	5.7	10869	758[1] 757[2]

(Western Red Cedar)

						← VBAR →		Volume per Acre	
Strat	Area (Acre)	Basal Mean	Area/Ac SE	Tree Tally	Trees	Mean	SE	Mean	SE
I	94	21.28	4.04	100	9	147.1	10.4	3130	634
II	85	29.42	4.56	125	12	118.5	14.7	3486	692
III	65	10.16	2.20	33	4	110.8	17.9	1126	304
Total	244	21.15	2.30	258	25	128.6	8.5	2720	353[1] 346[2]

(Western Hemlock)

						← VBAR →		Volume per Acre	
Strat	Area (Acre)	Basal Mean	Area/Ac SE	Tree Tally	Trees	Mean	SE	Mean	SE
I	94	6.38	1.55	30	8	202.0	19.2	1289	336
II	85	10.58	2.40	45	6	263.0	25.4	2783	686
III	65	1.23	0.97	4	2	203.3	9.0	250	197
Total	244	6.47	1.06	79	16	236.8	16.2	1533	277[1] 272[2]

(Combined Conifer Sawtimber)

						← VBAR →		Volume per Acre	
Spp	Area (Acre)	Basal Mean	Area/Ac SE	Tree Tally	Trees	Mean	SE	Mean	SE
DF	244	51.07	3.29	623	60	212.8	5.7	10869	758
WH	244	21.15	2.30	258	25	128.6	8.5	2720	353
RC	244	6.47	1.06	79	16	236.8	16.2	1533	277
Total	244	78.7	3.43	960	101	192.2	4.9	15122	764

---

[1] Standard error using stratification.

[2] Standard error by Goodman's product rule.

Tree Volume and Biomass

Moderator: J. Barrett

David Bruce and Timothy A. Max, Use of Profile in Tree Volume Estimation

Piermaria Corona and Agostino Ferrara, Developing a Simple Stem Taper Equation From Total Stem Volume Equations or Tables

Timothy G. Gregoire, George M. Furnival, and Harry T. Valentine, A Sampling Procedure for Board-Foot Volume of Standing Trees

Harry V. Wiant, Jr., and Geoffrey B. Wood, Estimating the Volume of a Radiata Pine Stand Using Importance Sampling

Andrew J.R. Gillespie, Sandra Brown, and Ariel E. Lugo, Biomass Estimates for Tropical Forests Based on Existing Inventory Data

Xiaoming Yang and Tiberius Cunia, Construction and Application of Biomass Regressions to Size-Classes of Trees

Boris Zeide, Fractal Geometry and Forest Measurements

Use of Profile Equations in Tree Volume Estimation

David Bruce and Timothy A. Max
Mensurationist and Station Statistician
Pacific Northwest Research Station, USDA, Forest Service
P.O. Box 3890, Portland, Oregon 97208-3890

ABSTRACT

Profile equations are used today for tree volume estimation because they are more flexible in estimating partial stem volumes and are just as accurate for total volume as two other systems. Profile models have been used for more than 200 years, but recent developments abetted by electronic computers make them more versatile today. Methods of accounting for variation in form, determining accurate actual volumes, and assuring desirable characteristics of models are described. Stem profile models currently used are discussed as well as anticipated improvements.

INTRODUCTION

Profile equations are generally used today as the basis for tree volume estimation because of their advantages over two other estimating systems: separate equations for total and partial volume of tree stems, and total stem volume estimators with secondary estimates of volume ratios for tips and stumps. Profile equations not only produce good estimates of total volume, but also give accurate estimates of both tip and stump volumes. This is not true of other systems where separate equations sometimes predict greater or lesser volumes for the sum of the parts of a stem than for the whole stem, and where ratio estimators require separate models for tips and stumps, if they are to be accurate. Flexibility, the ability to describe stem shapes accurately, is characteristic of profile equations but not of some older total volume estimators, which are neither accurate nor flexible and have only the virtue of simplicity. Profile equations can predict volumes to future utilization standards, but other systems usually require new data and analyses to do this. Based on the same data, these systems can be as accurate as the integrals of profile equations in estimating total volume, but not in estimates of partial stem volumes. This paper discusses some of what we know about realizing the advantages of profile equations and what is needed to make them even more useful.

The use of stem profiles is nearly as old as the history of tree volume estimation. In 1765, tree volumes were estimated by comparison to the volume of a cone with equal height and basal diameter (Oettelt 1765). In 1837, Smalian compared sections of tree stems to frustums of conoids with varying powers of exponents. All that is really new is the development of equations describing continuous curves that portray stem profiles more accurately than the older one-parameter models, and the use of definite integrals of these profile equations to estimate section volumes.

In 1987, a national committee was formed by the USDA Forest Service to develop a handbook on tree volume estimators. People from all Regions and some Stations who work with volume estimators prepared a draft of the handbook, which is scheduled for completion in 1990. The handbook includes detailed information related to the subject of this paper. Bruce helped write and revise some of its chapters, and discussions with other committee members helped shape and clarify ideas included in this paper.

VARIATION IN TREE FORM

Tree volume estimation would be simple if all trees had the same average shape. But even among species commonly used for major forest products, great diversity exists in branching, basal swell, and stem form. There are two related ways to measure stem form: form quotients or form classes are ratios of an upper stem diameter to basal diameter; form factors are ratios of stem volumes to volumes of cylinders with same height and basal diameter. These measures assume a simple conoid model for stem shape, and therefore, it is easy to write an equation for converting from one of these form measures to the other. Their high positive correlation makes what we say about one apply to the other. There is no single specification of the upper stem height at which diameter is measured, but basal diameter usually is measured at breast height (bh). Here the realities of field work outweigh theoretical advantages of using a relative height such as 5 or 10 percent.

Populations of trees can be stratified to accomodate major differences in form. Most populations are stratified by species or species group and by geographic location, and some are further stratified by age class and growth conditions--both site and cultural. Within these smaller strata, differences in form affect volume estimates for trees with the same height and diameter. A final stratification can be by size class, although regressions on size are often substituted for this.

Trees with unusual shapes are put in a special stratum or excluded from the sample and sometimes are given special rules for estimating volumes. For example, in most young-growth stands, trees with broken tops are unusual, but in old-growth stands trees whose tops have not broken or died back at some time are unusual. For young growth, special rules can be written to estimate volumes of trees with broken tops, but for old growth, many trees with minor breaks are included in the sampled population and no adjustment is needed. Old-growth trees with recent major breaks in upper stems require special volume estimates. Forked trees are often excluded from the sample. The definition of the population including strata descriptions, the identification of trees requiring special treatment in volume estimation, and the rules for rejecting trees from the sample are interrelated. All must be completely and carefully specified.

ACTUAL STEM VOLUME

Because true volume is not known, a definition of actual (or measured) volume is needed. It is not sufficient to say that actual volume is all the wood measured in the stem inside the bark from the ground line to the tree tip. It is necessary to describe how it is measured and how the actual volume is calculated. For example, diameter can be measured with a variety of instruments: a D-tape, calipers, a ruler, or an optical instrument. Measurements are made inside or outside the bark, or both, at specified intervals up the stem. The stem is assumed to be circular or elliptical, and allowances for convex deficits may be made. In the United States, wood volume is commonly estimated, although elsewhere many stem volume estimates are for wood and bark. When felled trees are measured, wood dimensions usually are measured directly, but in measuring standing trees, bark thickness is either measured or estimated. When the bark ratio at bh is used for the entire stem, large biases often result.

How many measurements are needed to determine actual volume and define stem shape? Useful, accurate estimates of total stem volume require total height, basal diameter, and one upper stem diameter; however, to test sufficiency or to calibrate such estimates, more measurements are necessary. Five regularly spaced measurements plus two extra near the base usually suffice. About twice this number are needed to define shape for a profile equation, although this should be tempered somewhat by tree size. If trees were all uniform in shape, fewer measurements would serve. But many trees have strange bumps and hollows on their profiles. The extra measurements help average out these deformities or overwhelm occasional wrongly recorded measurements. Extra measurements cannot correct for blunders in recording the independent variables: dbh, total height, or form quotient diameter. Nearby measurements can suggest errors in these records, but only top-quality field work will eliminate them. Breakdowns in copying, identifying, or storing accurate data can nullify such high-quality field work.

Given accurate data, the next step is to determine actual volume. Smalian's equation calculates the exact volume of paraboloids and has no notable bias for upper stem sections 8' or sometimes 16' long. Bias in the tip is negligible because the tip is small. Bias for butt sections, however, may exceed 15 percent (Bruce 1982). It is unfortunate that two studies comparing volumes computed with

Smalian's equation with those determined by water immersion failed to sort biased butt log volume estimates from unbiased upper log estimates (Young et al. 1967, Phillips and Taras 1987). Reporting only average bias gives the impression that Smalian's equation is biased for all logs. When measurements are close enough together (1' or 2') even the bias in butt sections may be small enough to ignore.

The old assumption that the stump is cylindrical may have been good enough for estimates of total stem volume, but should not be used when fitting stem profile equations. This assumption results in a false high bias for the estimate based on the integral of the stem profile equation because the estimate of actual volume was low.

In brief, actual stem volume is not true stem volume, but it is stem volume as actually determined. Carefully done, this determination is sufficient for practical purposes.

FITTING PROFILE EQUATIONS

The first step in data analysis is to convert measurements into relative heights and diameters (rh and rd). For example, rh is actual height of a measurement point, divided by total height. Relative height and diameter are dimensionless numbers with most of the high correlation between size and variation eliminated. Relative measures usually do not require weighting for valid regression analysis, and they are used to calculate form factors directly. Stem profiles based on these numbers often show gradual changes in tree shape with changes in tree size that can be used to produce accurate volume estimates. Unconverted measurements show large changes in volumes in trees of different sizes that mask these changes in shape.

Plotting scatter diagrams of relative heights and diameters is the best way to edit stem analysis data. Questionable data points stand out like sore thumbs, as do trees differing greatly in shape from others in the sample. Scatter diagrams of groups of trees in different classes suggest how stem form varies among these classes. Plotting first differences of rds can reveal aberrant pairs of successive measurements, as well as indicate location of join points for curve segments.

Form factors are useful for verifying total volume estimates. Being dimensionless, they change gradually with changes in tree size. Usually, except for very small trees, form factors within a height class decrease as diameters increase. Within diameter classes, form factors decrease slowly as height decreases from tall to average but then decrease more rapidly as height decreases further. This is largely due to inflation of dbh by butt swell in these shorter trees. There appears to be a lower limit of about one-third for form factor, with lower values for trees whose dbh is inflated by butt swell. Many studies in second-growth stands do not include butt swell problems because all the trees are too small.

Very small trees have special problems with measures of form. Dbh is a poor measurement to describe basal diameter for trees under 30' tall. Diameters of these very small trees should be measured at or near groundline. Often, their size is characterized by total height alone. On a 9' tall tree, bh is halfway to the tip and dbh is much smaller than diameter near the ground. This is an extreme example, but failure to recognize the problem has led to many very poor volume estimates for small trees. Evert (1969) suggested adjusting dbh to ground level by using a cone as a model. Bruce and DeMars (1974) used a hyperboloid as a model and substituted the adjusted diameter wherever dbh appeared in a total volume estimator. These adjustments are designed to correct overestimates of form factor by using dbh, but the same underestimates of basal diameter will inflate all estimates of relative diameters in small trees. Using height above bh as total height for such trees will inflate estimates of rh.

A few profile equations include changes in form related to differences in H/D ratio (total height divided by dbh), but a great many do not. Variation of tree shape within size class is so great that these changes are not easy to detect. When form quotient variation is studied, crown ratio often has the greatest effect and is accounted for by stratification or by inclusion as an independent variable (Dell 1979). This will hide the effect of H/D ratio, which often is highly correlated with crown ratio.

H/D ratio is always available when total height and dbh have been measured, while crown ratio requires extra measurements. H/D ratio reflects long-term growth conditions, but crown ratio may reflect only recent conditions. The effect of the H/D ratio on form may not be detected because of the narrow range in tree size included in many studies. Volume estimators that use single form factors may be valid for such limited data, but are poor general models and may provide unsatisfactory estimates for new data.

An example of a data set with a wide range of tree sizes was an optical-dendrometer stem analysis for old-growth redwood (Sequoia sempervirens (D.Don) Endl.). Tree size started at 40" and ended with two trees that were 220" dbh and about 310' tall. One of the 220" trees had about twice the volume of the other, so both records were rejected; the volume estimate ended at 180" where there were records for several trees. In such data sets, the importance of form factor is easy to see.

Generally in fitting profile equations, relative diameters should be computed by using basal diameter inside bark. Assmann (1961) said the well-known difference between the forms of Scots pine (Pinus sylvestris L.) and Norway spruce (Picea excelsa Link) applies only to their over-bark dimensions and that they are very similar in form after removal of the bark. Hilt (1980) used dbhib because bark thickness varies among oak species and appropriate bark-thickness factors can be used for each species. Bruce (1984) found for old-growth Sitka spruce (Picea sitchensis (Bong.) Carr.) and western hemlock (Tsuga heterophylla (Raf.) Sarg.) that using dbhib instead of dbhob produced polynomial profile equations with fewer terms and slightly lower standard errors.

In fitting profile equations, the usual measures of statistical significance in regression are not valid because of lack of independence among observations from the same tree, and to a lesser extent among trees within even-aged stands. The estimate of true variance of the former falls between the mean square calculated by using degrees of freedom based on number of measurement points and one based on number of trees measured. Because the correct degrees of freedom will not differ among models, the calculated standard errors are good indicators of relative precision of models under study.

DESIRABLE CHARACTERISTICS OF STEM PROFILE EQUATIONS

Consideration of desirable characteristics helps select models to use for stem profile equations. Precision and accuracy are first, as they should be for most estimators. Achieving precision and accuracy requires precise definition of the population, faithful adherence to the definition in drawing the sample, and good technique in measuring the trees and recording, storing, and processing data. These requirements are not peculiar to profile equations.

Profile equations are used to estimate both stem diameters and volumes, so it is desirable to have minimum-variance, unbiased estimates of both d (upper-stem diameter ib) and d^2. Estimates of d^2 are used in making volume estimates and estimates of d in estimating the scale of different products in standing trees. No single equation will produce such unbiased estimates, so compromises or adjustments are needed. One compromise is to fit two equations and minimize the pooled error. This can be done by using a standardized loss function (Reed and Green 1984). This may not be appropriate if there is a compelling reason for maximizing the precision of one of the estimates. It is unnecessary to calculate the error in the volume estimate in making this compromise because volume is obtained by integrating sectional area between stated heights.

Another solution to the d and d^2 problem was used by Kilkki et al. (1978). They fit d by using least squares and adjusted estimates of d to those that might have been derived from fitting d^2 by adding the correct amount to get the quadratic mean. This makes the needed adjustment to the total of their transformed estimates. But this solution was not correct for some equations recently fitted to stem analysis data. The easiest way to determine the correct adjustment seems to be to fit equations in d and d^2, and once this is done the adjustment is no longer needed. A general adjustment rule should be sought, however, so that a single equation can be fitted, with little bias in the adjusted estimates of the unfitted variable.

The ability to rewrite the equation to express relative height as a function of relative diameter is often considered desirable. This manipulation is possible for many profile equations, but the coefficients so derived are not the same as those obtained by fitting the inverse function by least squares. When this algebraic manipulation is possible, an equation can be fitted that accounts for error in both variables.

Because the profile equation is mainly used for volume estimation, it should be integrable in closed form. If a profile equation requires numerical integration, it should be outstanding in all other respects to be acceptable.

Estimators should make logically correct estimates at specific points on the stem, such as zero volume at zero height, zero diameter at tip of tree and nowhere else, and estimates equaling the actual values of the independent variables. All these can be accomplished by conditioning the equations.

The final desirable characteristic we mention is flexiblity. By flexibility, we mean ability to describe a great variety of stem forms. This includes reflecting the average effect of H/D ratio and crown ratio on form factor, as well as describing differences in butt swell caused by sloping terrain and changes in rooting characteristics caused by soil texture or moisture. Flexibility is obtained by using functions of factors that control form as exponents of rh or as coefficients of terms in the profile equation. Many equations include D and H merely to convert form factors to volumes.

STEM PROFILE MODELS

We believe the only equations worth considering for stem profiles are those whose independent variables are terms raised to whole or fractional powers (power functions), including simple and segmented polynomials. This rules out logarithmic or trigonometric curves, Behre hyperbolas, and a few other exotic curves that have been used to describe stem profiles. A third-degree polynomial can be fit to 11 equidistant points on Behre hyperbolas that have coefficients between 0.3 and 0.5, and the numerical values of all but one estimate will be within 1 percent of the exact values of the hyperbola. Because the error in fitting curves to stem profile data is between 5 and 10 percent, any claimed advantage of the hyperbola is an illusion. Similar exercises probably could be carried out with all other curve forms. Power functions have simple closed-form definite integrals, which is not true of some other curve forms.

Typical stem profile curves for rd have an inflection point at about 20 percent of height. At this point, the curvature changes from concave to convex. This is not a change that is visible in the forest, but it shows up on graphs having vertical scales compressed to between 1/50 and 1/100 of the horizontal scale. It may disappear if rd^2 instead of rd is plotted against rh. The basal concavity still shows, but the upper stem usually is transformed to a straight line or a slightly concave curve. The point of tangency to the straight line, if it exists, may be at or slightly below the former inflection point. This distinction is mentioned because the rd^2 transformation is common, as is the search for the height of the inflection point.

Many good stem profile curves are segmented polynomials (Max and Burkhart 1976). The lower join point is at or between bh and the inflection point. The segmentation is needed because butt swell can be described by terms with coefficients or exponents differing from those describing the upper stem. There are two reasons for this: first are physical laws affecting stem shape. Butt swell is caused by horizontal roots joining a vertical stem, and shape of the upper stem is largely determined by requirements of mechanical support of the crown on a vertical pole. Second, the distance from bh to top of stump usually is a fixed amount, while the important distances of the upper stem are all calculated relative to the total height of the tree. This gives a wide range of relative heights at stump height. In small trees, use of diameter at bh to form rd causes greater variation in rd below than above bh.

A problem with most segmented curves is that they need to be kept smooth and continuous at the join point. They are made continuous by conditioning submodels to be connected at these points and smoothed by conditioning their first and higher order derivatives (Max and Burkhart 1976). Conditioning the first derivative usually smoothes the curve adequately. Another solution to smoothing a segmented profile curve has been to shift the origin to the join point and fit separate power functions above and below this point (Demaerschalk and Kozak 1977). These curves are continuous and have the minor flaw that their tangents are vertical at the origin where they join. The amount the axis should be tilted might be about one-third the amount indicated by the H/D ratio, and this possibly can be ignored.

High-degree polynomials with the origin at the tip of the tree can describe the extreme concavity at the base of the tree (Bruce et al. 1968). They need no conditioning to ensure smoothness because the value of rh raised to a high power essentially disappears when rh is a little smaller than 1.0. They have a problem common to many polynomials that describe curves with maxima or inflections; that is, the coefficients of the terms describing these shapes have opposite signs. This works well when the coefficients are fixed values but can lead to strange curves when the coefficients are functions of tree size or other factors that affect stem form.

Equations including rh raised to powers that are functions of rh describe shape well and seem to be flexible enough to demonstrate changes in form factor (Ek and Kaltenberg 1975, Kozak 1988). They unfortunately require numerical integration.

Where information about form quotient is available, it should be used directly by forcing the curve through the upper stem diameter (Czaplewski and McClure 1988). It may be difficult to combine this with the shift in origin that makes it easier to describe butt swell. The appropriate mean form quotient to use with this equation may be hard to determine when only height and basal diameter are measured.

An unpublished curve developed many years ago by Bruce was a polynomial including powers of both relative height and the complement of relative height. The terms with zero at tree base describe the upper stem with little effect on the shape of the lower stem; terms with zero at tip can describe the butt swell. This model deserves further research to see if it solves old problems with stem profile equations, without introducing new ones.

The polynomial curve that Bruce et al. used in 1968 had the dependent variable written as (y^2-$x^{3/2}$). With both x and y in ratio measure (ranging in value from 0 to 1), this combination is the difference between an observed and an estimated dependendent variable. This form could be pushed a little further by using a slightly more complicated model that includes butt swell. Deviations from this model could be analyzed as functions of relative height and other variables that cause changes in shape.

None of the curve types discussed so far directly solves the problem of using relative height above bh and actual height below bh. This might best be done with two equations both estimating rd and having vertical tangents at a join point near bh. The upper-stem equation would have rh as an independent variable and the lower-stem equation would have actual height.

EVALUATION

A favorite method used to select mathematical models and evaluate volume estimators is data splitting. This often is done with little regard for the methods recommended for evaluating industrial production data by one of the first people to describe it (Snee 1977). Instead of splits that are used to expose differences within the data set, the splits seem to be attempts to ensure similarity in the two parts. When it is not possible to collect new data for validation, Snee suggests using time as a basis for splitting. Many of his comments on data splitting are applicable to tree measurement data. His analysis does not include large data sets prepared by different crews in different parts of a large region. With such data, analysis of covariance is a logical extension of his methods and would be applicable as long as subsets have comparable variation. This not only validates the estimator, but also gives an indication of the average deviations from the model to be expected from future samples measured by other crews in other parts of the region. It further provides some information about whether all samples seem to be drawn from the same population. Selection of robust models over those showing large differences among coefficients fitted in different subsets is preferred.

Stem profile equations rely on integration to make volume estimates; therefore, a natural evaluation is to compare direct estimates of total stem volume with these integrals. The correct way to proceed is to compare each of the estimates with the numbers recorded as actual volumes. Deviations, actual minus estimated over estimated, (a-e)/e, should show no correlation with the independent variables height or dbh, and the estimate with the lower standard error can be judged the more precise. Nonsignificant regressions of volume differences over dbh for spruce and hemlock had standard errors of 0.1085 for logarithmic total volume estimates and 0.1074 for integrals of profile equations (Bruce 1984). With no adjustment for the logarithmic transformation, these were close enough to call the same, even if the profile equation had a slight advantage.

CONCLUSIONS

If reasonable care is taken in gathering data and fitting profile equations, the equations should produce volume estimates at least as accurate as any total volume estimator based on the same data. Because they require more detailed measurements than total volume estimates, they are likely to be more accurate than many older volume estimators. Apparent biases in the volume estimates often are found to be biases in the initial estimates of actual volume for individual trees, which were tolerable with total volume estimates but not with profile equations.

We see three major ways that profile equations will be improved in the next few years to make them even more versatile and useful:

(1) There will be better ways to handle the problem of reducing bias in estimates of both d and d^2.

(2) New ways will be developed to reduce the problems introduced by measuring basal diameter at bh, which underestimates basal diameters of small trees and overestimates those of large trees because of butt swell. Methods also will be developed to handle inflated estimates of rh at stump height for small trees.

(3) Triple entry estimators, using dbh, total height, and an upper-stem diameter, will be developed that increase accuracy of volume estimates wherever it is feasible to measure an upper stem diameter on standing trees.

REFERENCES

Assmann, E., 1961: Waldertragskunde. BLM Verlagsgesellschaft, München, 490 pp.

Bruce, D., 1982: Butt-log volume estimators. Forest Science, vol. 28, pp. 489-503.

Bruce, D., 1984: Volume estimators for Sitka spruce and western hemlock in coastal Alaska. pp. 96-102 in Inventorying forest and other vegetation of the high latitude and high altitude regions, Proceedings, International Symposium, Fairbanks, Alaska.

Bruce, D., Curtis, R.O. and VanCoevering, C., 1968: Development of a system of taper and volume tables for red alder. Forest Science, vol. 14, pp. 339-350.

Bruce, D. and DeMars, D.J., 1974: Volume equations for second-growth Douglas-fir. USDA Forest Service Research Note, PNW-239, 5 pp.

Czaplewski, R.L. and McClure, J.P., 1988: Conditioning a segmented stem profile model for two diameter measurements. Forest Science, vol. 34, pp. 512-522.

Dell, T.R., 1979: Potential of using crown ratio in predicting product yield. pp. 843-851 in Forest resource inventories, W.E. Frayer, ed. Colorado State University, Fort Collins.

Demaerschalk, J.P. and Kozak, A., 1977. The whole-bole system: a conditioned dual-equation system for precise prediction of tree profiles. Canadian Journal of Forest Research, vol. 7, pp. 488-497.

Ek, A.R. and Kaltenberg, M.C., 1975: Generalizaion of tree taper and volume models. Abstract, Forest Mensuration Meeting, October 1975, Lake Barkley Lodge, Kentucky, 2 pp.

Evert, F., 1969: Use of form factor in tree-volume estimation. Journal of Forestry, vol. 67, pp. 126-128.

Hilt, D.E., 1980: Taper-based system for estimating stem volume of upland oaks. USDA Forest

Service Research Paper NE-458, 12 pp.

Kilkki, P., Saramaki, M. and Varmola, M., 1978: A simultaneous equation model to determine taper curve. Silva Fennica, vol. 12, pp. 120-125.

Kozak A., 1988: A variable-exponent taper equation. Canadian Journal of Forest Research, vol. 18, pp. 1363-1368.

Max, T.A. and Burkhart, H.E., 1976: Segmented polynomial regression applied to taper equations. Forest Science, vol. 22, pp. 283-289.

Oettelt, 1765: "Praktischer beweis, das die mathesis bey dem forstwesen unentbehrliche dienste thue". Eisenach. (not seen)

Phillips, D.R. and Taras, M.A., 1987: Accuracy of log volume estimates by density and formulas compared with water displacement. Forest Products Journal, vol. 37, pp. 37-42.

Reed, D.D. and Green, E.J., 1984: Compatible stem taper and volume ratio equations. Forest Science, vol. 30, pp. 977-990.

Smalian, L., 1837: Beitrag zur holzmesskunst. Stralsund. 124 pp.

Snee, R.D., 1977: Validation of regression models: methods and examples. Technometrics, vol. 19, pp. 415-425.

Young, H.E., Robbins, W.C., and Wilson, S., 1967: Errors in volume determination of primary forest products. pp. 546-562 in Proceedings, 14th Congress, International Union of Forest Research Organizations, Part VI, Section 25, pp. 546-562. Munich.

Developing a Simple Stem Taper Equation from Total Stem Volume Equations or Tables

Piermaria Corona
Head of the Silvicultural Section (1)

Agostino Ferrara
Head of the Forest Inventory and Land Survey Service (1)

(1) S.A.F. - Centro di Sperimentazione Agricola e Forestale
via di Casalotti 300, 00166 Rome, Italy

Authors are also, respectively, Adjunct Professor of Silviculture and
Adjunct Professor of Forest Inventory at Basilicata University - Italy

ABSTRACT

In forest inventory, tree stems, chiefly those of conifers, can be assimilated, in a first approximation, to paraboloids of revolution generated from simple taper equations, like the Ormerod's one:

$$d = D^*((H-h)/(H-13))^b$$

where d = stem diameter at h height; H = total stem height; D = diameter at breast height (i.e., diameter at 13dm-height); b = function's parameter.

This taper equation does not prove fully correct from a theoretical point of view, due to the fact that tree stems are usually regarded as solids of form sectionable into truncated paraboloids of revolution, each characterized by a different exponent.

In practice, however, equations like the Ormerod's one, with a constant exponent for the whole stem profile, have found a fairly good applicability in forest inventory, since, despite their simple structure, they can provide rather acceptable approximations in merchantable volumes estimates.

The applicative profit of using a taper equation like the Ormerod's one is demonstrated by the fact that the b exponent can be calculated from already existing total stem volume equations or tables (even if elaborated by graphical smoothing), through the procedure developed by the Authors and illustrated in the paper.

The procedure is based on an iterative algorithm using the stem form factor determined for each hypsodiametrical class from the selected volume equation or volume table. The compatible system to predict merchantable volumes is reported.

The performance of such estimation procedure, evaluated using stem analysis data on Monterey pine from Italian plantations and on Scotch pine from stands of north-eastern Italy, proved satisfactory and suitable for inventory purposes.

INTRODUCTION

In forest inventory, tree stems, chiefly those of conifers, can be assimilated in a first approximation to paraboloids of revolution generated from taper curves expressed by the general function:

$$d/2 = a*(H-h)^{b} \qquad [1]$$

where d = stem diameter at h height; H = total stem height; a, b = function's parameters.

Having $2*a*(H-13)^{b}$ = diameter at breast height (i.e., diameter at 13dm-height = D), Ormerod's taper equation (1973) can be derived from the above general formula:

$$d = D*((H-h)/(H-13))^{b} \qquad [2]$$

The exponent b is directly conditioned by the general stem form: in the case of the cone, b takes on a value of 1; in the case of the semicubic paraboloid, a value of 1/3. Actually, when considering the real general stem forms, it can be observed that the b value usually fluctuates between 0.5 and 0.9, according to the species and, for each species within a given geographical area, according to the hypsodiametrical class of each stem.

The Ormerod's taper equation does not prove fully correct from a theoretical point of view, due to the fact that tree stems are usually regarded as solids of form sectionable into truncated parabolids of revolution, each characterized by a different exponent. In practice, however, equations as above [2], with a constant exponent for the whole stem profile, have found a fairly good applicability in forest inventory, since, despite their simple structure, they can provide rather acceptable approximations in merchantable volumes estimates (Reed and Byrne 1985).

The applicative profit of using a taper equation like the Ormerod's one is mainly demonstrated by the fact that the b exponent can be directly calculated from already existing total stem volume equations or tables (even if elaborated by graphical smoothing), through the following procedure developed by the Authors, a preliminary version of which has already been published (Corona and Ferrara 1989).

THE PROCEDURE DEVELOPMENT

By rotating equation [1] around the axis of equation $d/2 = 0$ between $h = 0$ and $h = H$, paraboloids' volume (V) general formula is obtained:

$$V = 1/(2*b+1)*PI*a^{2}*H^{(2*b+1)} \qquad [3]$$

from which, being $PI = 3.141593$ and $a*H^{b}$ equal to the radius of section at height $h = 0$, it can be deduced that the total stem volume can be regarded as the product of a cylinder of section $PI*a^{2}*H^{(2*b)}$ and height H, multiplied by a coefficient of reduction, commonly known in Europe as geometric form factor (GF), equal to:

$$GF = 1/(2*b+1) \qquad [4]$$

Usually, for greater convenience and precision, the area of the stem base section is not taken into account, measuring on the contrary that of b.h. section. Therefore, in volume estimations, GF cannot be used, but it is necessary to resort to another coefficient of reduction, currently known as (ordinary) form factor (*FF*), resulting from the ratio between the total stem volume and the volume of a cylinder with the same stem height (*H*) and with the section equalling that of the stem at breast height, i.e. - in the same terms as equation [1] - $PI*a^2 (H-13)^{(2*b)}*H$.

Expressing the paraboloids' volume general formula as a function of *FF*, the equation obtained is the following:

$$V = FF*PI*a^2 (H-13)^{(2*b)}*H \qquad [5]$$

The latter, compared with [3] , allows the calculation of *GF* as a function of *FF* and *b*:

$$GF = FF*((H-13)/H)^{(2*b)} \qquad [6]$$

For calculating directly *GF* as a function of *V*, besides *b*, it is only necessary to replace *FF* with $V/(PI/4*D^2*H)$.

Together with the equation [6] , keeping into account also equation [4], which can be converted into the expression

$$b = (1/GF-1)/2 \qquad [7]$$

it can be observed that, for evaluating *b*, the only necessary element, besides *H*, proves *FF* (or *V*), which can be directly assessed for every hypsodiametrical class from either the volume equation or volume table.

On the basis of equation [6] and [7], *b* can be calculated by an iterative algorithm, according to the following procedure:

1) an approximate value of *b* is introduced into [6] and the corresponding *GF* is calculated through the value of *FF* (or *V*) deduced from the given volume equation or volume table;

2) the *GF* value calculated with [6] is introduced into [7], so as to obtain a new value of *b*;

3) if such value coincides with that introduced into [6] the calculation is ended; otherwise, the value of *b* calculated at point 2) is used to estimate a new value of *GF* through [6]; then, on the basis of such value, another *b* is calculated through equation [7] and so forth, till the value of *b* introduced into [6] and that calculated by [7] coincide (with an approximation that should not exceed the sixth decimal figure).

The applicative importance of the proposed procedure lies in the fact that, once that *b* is calculated for a given hypsodiametrical class, it is possible to determine the volumes to any height or upper stem diameter from the stems belonging to that class.

In fact, the stem volume to any height or upper stem diameter (*Vx*) is calculable according to the expression:

$$Vx = Rx*V \qquad [8]$$

where the total stem volume (*V*) is established by volume equation or table and *Rx* coefficient is easily estimable as a function of either height *h*

$$Rh = 1-((H-h)/H)^{(2*b+1)} \qquad [9]$$

or upper stem diameter *d*

$$Rd = 1-((H-13)/H)^{(2*b+1)}*(d/D)^{((2*b+1)/b)} \qquad [10]$$

given *D*, *H* and known the value of *b* by the iterative algorithm.

To evaluate merchantable volumes, it suffices to estimate the volumes to both heights or upper stem diameters determining the desidered wod assortments along the profile and to calculate their volumes by difference.

TESTING AND CONCLUSIONS

The proposed procedure was tested on the basis of the following total stem volume equations worked out by the multiple regression techniques currently used in forest mensuration and concerning, respectively, the Monterey pine grown in Italian plantations and the Scotch pine from north-eastern Italy:

Monterey pine - $V=8.902+4.435*D^2+0.305*D^2*H$ (SEE = 28 dm^3)

Scotch pine - $V=1.516*D^{2.034}*H^{0.726}$ (SEE = 74 dm^3)

where *V* = total stem volume (in cubic decimeters); *D* = b.h. diameter (in decimeters); *H* = total stem height (in decimeters).

The suggested procedure was tested on the inventory data obtained from the stem analysis carried out on 1902 Monterey pines (data ranges: *D* = 0.9-4.0 dm; *H* = 70-240 dm; *V* = 30-1250 dm^3) and 112 Scotch pines (data ranges: *D* = 1.3-4.9 dm; *H* = 70-275 dm; *V* = 60-2110 dm^3), previously sampled for the elaboration of the above mentioned volume equations.

For every hypsodiametrical class, computation of exponent *b* was performed after the program reported in table 1.

The diameters' estimate along the stem profile proves satisfactory, with absolute percent residuals averaging 6% for Monterey pine and 6.5% for Scotch pine. The percent of variation explained attains high levels for both species: respectively, 97% and 95%. It should be stressed that, owing to the simple structure of the mathematical model, equation [2] is unable to follow exactly the profile along the whole stem: there is, in fact, a systematic tendency to positive errors towards the bottom and the top and to negative errors in the middle of the stems (Table 2).

To test the reliability of such procedure, the observed volume values of 2m-long logs taken from the sampled trees at different heights were compared with those predicted by equation [8]. On the whole, the average absolute residual proves to be 10% for Monterey pine and 12% for Scotch pine. The percent of variation explained attains high levels, equalling 98% and 94%, respectively, for the two species. The predicted log volumes suffer the systematic errors observed for diameters: however, the errors prove inconsistent from an operative standpoint (Table 3).

The suggested procedure appears suitable enough for applicative purposes. Its use can be easily generalized to those species for which, within a geographical area, total stem volume equations or tables are available, so as to meet the commonest mensurational needs in forest inventory.

Table 1. A simple Basic program to perform the computation of *b*.

```
10  '------------------------------------------------------------------------------------------
20  '      Calculation of Ormerod's taper equation exponent b
30  '------------------------------------------------------------------------------------------
40  '
50  INPUT "Diameter at breast height        (dm)              = "; D
60  INPUT " Total stem height               (dm)              = "; H
70  INPUT "Total stem volume                (dm3)             = "; V
80  PI= 3.141593
90  '
100 ER= 0.000001                                    ' level of approximation
110 FF= V/(PI/4*D^2*H)                              ' form factor
120 B1= 0.7                                         ' approximate starting b value
130 '
140 GF= FF*((H-13)/H)^(2*B1)                        '⎤
150 B= (1/GF-1)/2                                   '  iterations
160 IF ABS(B-B1)>ER THEN B1=B:GOTO 140              '⎦
170 '
180 PRINT "b = "; B
190 '
200 END '---------------------------------------------------------------------------------------
```

Table 2. Performance of the taper equation [2] by location in tree stem.

	Monterey pine			Scotch pine		
proportion of total height	no. of obs.	AAR	AR	no. of obs.	AAR	AR
$0.0 \leq h/H < 0.1$	5261	0.12	+0.10	231	0.05	+0.03
$0.1 \leq h/H < 0.2$	3632	0.06	-0.04	133	0.12	-0.12
$0.2 \leq h/H < 0.3$	2516	0.09	-0.05	124	0.12	-0.10
$0.3 \leq h/H < 0.4$	2525	0.08	-0.03	131	0.11	-0.06
$0.4 \leq h/H < 0.5$	2480	0.09	+0.00	138	0.12	+0.02
$0.5 \leq h/H < 0.6$	2564	0.10	+0.02	120	0.15	+0.08
$0.6 \leq h/H < 0.7$	2569	0.11	+0.02	126	0.21	+0.15
$0.7 \leq h/H < 1$	6749	0.09	+0.02	206	0.18	+0.15

$AAR = \Sigma|o_i - p_i|/N$

$AR = \Sigma(o_i - p_i)/N$

where: o_i = observed diameter (in dm); p_i = predicted diameter (in dm); N = number of observations.

Table 3. Performance of volume estimation system by location in tree stem.

	Monterey pine			Scotch pine		
proportion of total height	no. of obs.	AAR	AR	no. of obs.	AAR	AR
$0.0 \leq h/H < 0.1$	1902	4.2	+2.3	114	6.5	+5.1
$0.1 \leq h/H < 0.2$	1902	4.4	-2.8	114	11.8	-10.9
$0.2 \leq h/H < 0.3$	1896	4.1	-1.6	114	10.5	-8.7
$0.3 \leq h/H < 0.4$	1806	3.3	-0.5	112	8.1	-4.3
$0.4 \leq h/H < 0.5$	1520	3.2	+0.5	109	7.5	+1.2
$0.5 \leq h/H < 0.6$	1122	3.5	+1.1	103	8.9	+6.0
$0.6 \leq h/H < 0.7$	793	3.6	+1.7	89	9.9	+7.8
$0.7 \leq h/H < 1$	519	3.4	+2.0	68	11.6	+9.2

$AAR = \Sigma|o_i - p_i|/N$

$AR = \Sigma(o_i - p_i)/N$

where: o_i = 2m-long log observed volume (in cubic decimeters); p_i = 2m-long log predicted volume (in cubic decimeters); N = number of observations.

ACKNOWLEDGEMENTS

The Authors would like to thank H.E. Burkhart, B. Herrigl and J.A. Scrivani for thoughful comments.

REFERENCES

Corona, P. and Ferrara, A., 1989: A simple stem taper equation from total stem volume equations or tables. The Compiler, vol. 7, no. 1, pp. 4-8

Ormerod, D.W., 1973: A simple bole model. Forestry Chronicle, vol. 49, no. 3, pp. 136-138.

Reed, D.D. and Byrne, J.L., 1985: A simple variable from volume estimation system. Forestry Chronicle, vol. 61, no. 2, pp. 87-90.

A Sampling Procedure for Board-foot Volume of Standing Trees

Timothy G. Gregoire
Assistant Professor of Forest Biometrics
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

George M. Furnival
J. P. Weyerhaeuser, Jr., Professor of Forest Management
School of Forestry and Environmental Studies, Yale University
New Haven, Connecticut 06511

Harry T. Valentine
Research Forester
Northeastern Forest Experiment Station, USDA Forest Service
Hamden, Connecticut 06514

ABSTRACT

A sampling procedure is presented which uses a taper function to provide approximate volumes in board feet of the logs in a standing tree. Logs are then sampled with probability proportional to approximate volume in order to estimate the board-foot volume of the tree. Variations in the determination of the number of logs in the tree and in the scaling procedures applied to selected logs are discussed. The performance of the sampling procedure is examined when it is applied to 223 yellow-poplar trees whose actual volumes are known from stem analysis information.

INTRODUCTION

The determination of sawtimber tree volume in units of board feet is prevalent in the United States. Tables and equations have been prepared for this purpose with the common objective of providing an estimate of the board-foot lumber yield of a standing tree that corresponds closely to the volume obtained by felling the tree, bucking it into logs, and scaling its logs (Avery and Burkhart 1983). Of these, the Mesavage-Girard form-class tables are perhaps the most widely known. Mawson and Connors (1987) smoothed the International 1/4 inch (I-1/4) tabular values by regression. Many other equations have been developed with utilization study data. Recent work includes that of Hann and Bare (1978) for eight species of pine, fir, and aspen native to the southwestern U.S. region; Saucier, Phillips, and Williams (1981) and Burkhart *et al.* (1987) for the major pine species in the southeastern region; and Burk, Hans, and Wharton (1989) for the 16 species groups native to the northeastern region.

As an alternative to volume estimation based on a fitted equation or tabular presentation, a sampling procedure can be applied. A sampling approach may be preferable when a volume equation lacks desired accuracy or when it is desired to check the performance of an equation whose accuracy is suspect. We present a method of probability sampling an individual tree for the purpose of estimating its board-foot volume (BFV) as it would be scaled on the felled and bucked tree.

SAMPLING PROCEDURE

In brief, the procedure consists of partitioning the stem of the standing tree into logs, selecting one or more logs at random with probability proportional to predicted BFV, scaling each selected log, and inflating each scaled log's volume by the per-draw selection probability of the log. Tree volume is estimated as the sample average of the inflated log volumes. In the following, we discuss details of each of the above steps and then conclude with an examination of the performance of various estimators of BFV when applied to a collection of yellow poplar (*Liriodendron tulipifera* L.) trees.

Partition of Stem into Logs

One straightforward approach involves measuring the merchantable height of the tree and partitioning the stem into successive 16-foot logs, allowing for the possibility of a shorter top log. Variations on this theme include the use of a standard log length other that 16 feet and the partition of the stem into non-contiguous logs. The latter strategy may arise if each stem is visually inspected for defect and defective portions of the stem are excluded from scaling.

If merchantable height is not measured, a taper function can be used to determine merchantable height based on dbh and the minimum scaling diameter. With this approach, it seems most sensible to partition the stem into successive logs of some standard length. It is possible that the taper function will determine more of fewer logs than actually exist. The latter case is a source of bias since part of the merchantable portion of the tree is given zero chance of inclusion into the sample. No bias is introduced in the former case if the extraneous logs are ultimately scaled to have zero BFV, but the precision of the volume estimator is likely to be less than it would be with a taper function that correctly determines the number of logs.

Selection of Logs

Regardless whether merchantable height is measured or is determined indirectly through use of a taper function, a taper function is used in this step to provide approximate scaling diameters of the logs in the tree stem. In the demonstration example that we present, we derived the approximate scaling diameter by applying an inside:outside-bark diameter ratio to the outside-bark diameter predicted by the taper function. Alternatively, one may choose to use either a more elaborate adjustment (cf. Hilt, Rast, and Bailey 1983) or an inside-bark taper function.

The approximate scaling diameters are used with a log rule to determine the approximate log volume of the logs. The approximate log volumes then serve as the basis for probability proportional to size (pps) sampling of one or more logs. The closer that these approximate scaled volumes are to the actual log volumes, the more precise will be the estimate of board-foot volume in the tree. Any log rule can be applied, although a formula-based rule is reasonable because it can be codified into a program for a hand-held calculator or field computer.

To avoid the introduction of a selection bias, the approximate BFV of each log should be determined regardless whether the approximate scaling diameter exceeds the minimum scaling diameter considered merchantable. Without this provision, zero probability of selection may be assigned to one or more of the upper logs in the tree, thereby biasing the volume estimator downward. Conversely, the precision of the estimator may be improved if these logs are excluded from the sampling frame. Even with this provision, care must be taken to avoid selection bias. For example, we used the I-1/4 rule, which for a sufficiently small scaling diameter will provide a negative BFV. In this event, the approximate log volume must be re-assigned a non-negative value so that the corresponding selection probability of the log is non-negative.

Logs may be selected with or without replacement. Due to the small number of logs in a tree, a with-replacement strategy may be quite inefficient compared to a without-replacement strategy. We suggest, however, that an estimate that is sufficiently precise for purposes of most inventories may be obtained by selecting a single log per tree, especially if this selection is in the second stage of a two-stage sampling effort.

Log Scaling

The scaling diameter of each selected log must be determined in order to scale the log. It is not likely that scaling diameter can be measured on a standing tree in a manner that generally will coincide with that obtained with a scale stick applied to a log on the ground. In this regard we have in mind the use of inside-bark diameter (DIB) as a substitute for scaling diameter. DIB may be computed from a measure of outside-bark diameter (DOB) minus twice the measured thickness of bark. Because this level of measurement may not be feasible or worthwhile, alternatively one may adjust the measured DOB by some factor in order to derive DIB. Because scaling diameters are rounded to the nearest inch, the need for accuracy in the measurement of DIB is somewhat muted.

Estimator of Tree Volume

The BFV of a tree is an inherently ambiguous quantity because it depends on the manner in which the tree is bucked among many other factors (Grosenbaugh 1964). We define the target value to be the aggregate scaled volume of the logs that partitioned the tree stem in the earlier step, plus any additional top logs that were not included in the partition, but that exceed the minimum scaling diameter. Let Y_i denote this target value for the ith tree in the forest, $i = 1, 2, \ldots, N$. Let m denote the number of logs (not necessarily distinct) that are selected from the ith tree. The estimator of Y_i is:

$$\widehat{Y}_i = \frac{1}{m} \sum_{k=1}^{m} \widehat{Y}_{i(k)}$$

where $\widehat{Y}_{i(k)}$ is the estimator of tree volume based on the kth selected log:

$$\widehat{Y}_{i(k)} = Z_{ij} / Q_{ij}$$

Z_{ij} is the volume of the jth log as scaled on the standing tree, and Q_{ij} is the selection probability of the jth log used in the list sampling. Defining $t_{ijk} = 1$ if the jth log of the ith tree is chosen on the kth draw, or $t_{ijk} = 0$ otherwise, then $\widehat{Y}_i$ can be re-expressed as:

$$\widehat{Y}_i = \frac{1}{m} \sum_{k=1}^{m} \sum_{j=1}^{M_i} \frac{Z_{ij} t_{ijk}}{Q_{ij}}$$

where M_i denotes the number of logs in the ith tree. If selection is with replacement, the variance of $\widehat{Y}_i$ is $\mathrm{Var}(\widehat{Y}_i) = \sigma_i^2 / m$, where

$$\sigma_i^2 = \sum_{j=1}^{M_i} Q_{ij} \left(\frac{Z_{ij}}{Q_{ij}} - Z_i \right)^2 \qquad \text{and} \qquad Z_i = \sum_{j=1}^{M_i} Z_{ij}.$$

An estimator of $\mathrm{Var}(\widehat{Y}_i)$ is:

$$\mathrm{var}(\widehat{Y}_i) = \frac{1}{m(m-1)} \sum_{k=1}^{m} (\widehat{Y}_{i(k)} - \widehat{Y}_i)^2$$

Two-stage Estimator of Total Volume

Given N trees, the target parameter is:

$$Y = \sum_{i=1}^{N} Y_i$$

Let n trees be selected in a first stage of sampling and let P_1 denote the selection probability of the ith tree in the first-stage sample. For each tree selected, BFV is estimated by the procedure outlined above. An estimator, $\widehat{Y}$, of total volume is the average of values $\widehat{Y}_i/P_i$ over the n (not necessarily distinct) trees chosen in the first stage. If sampling with replacement is used, and the second-stage sample size, m, is constant over first-stage selections, then

$$\widehat{Y} = \frac{1}{n} \sum_{h=1}^{n} \sum_{i=1}^{N} \frac{\widehat{Y}_i s_{ih}}{P_i}$$

where $s_{ih} = 1$, if the ith tree is selected on the hth draw in the first stage, or $s_{ih} = 0$, otherwise.

If with-replacement sampling is conducted at both stages, the mean square error of $\widehat{Y}$ is:

$$\mathrm{MSE}(\widehat{Y}) = \frac{\sigma_B^2}{n} + \frac{\sigma_W^2}{nm} + (Z - Y)^2 \qquad [1]$$

where

$$\sigma_B^2 = \sum_{i=1}^{N} P_i \left(\frac{Z_i}{P_i} - Z \right)^2$$

is the between-tree variance,

$$\sigma_W^2 = \sum_{i=1}^{N} \sigma_i^2 / P_i$$

is the pooled within-tree variance, and

$$Z = \sum_{i=1}^{N} Z_i .$$

EXAMPLE

Detailed stem measurements of 223 yellow poplar trees provided data which enabled us to apply the sampling procedure. These trees were the subset of 336 trees measured by Beck (1963)[1] that contained at least one 8-foot log. Among the 223 trees, dbh (in inches) ranged from 9.8 – 30, with a mean of 16.7; total height, H_T (in feet), ranged from 63 – 138, with a mean of 105. Stem DOB and bark thickness were measured at intervals of 4 feet above a 1 foot stump. Each tree was partitioned into successive 16-foot logs and a top log of either 8, 12, or 16 feet. The volume for each log was computed according to the I-1/4 rule using DIB at the top of each log as the scaling diameter of the log and a minimum scaling diameter of 8 inches. The BFV of a tree was reckoned as the accumulated board feet of logs in the tree. Among the 223 trees, BFV ranged from 22 – 1497, with a mean of 373 and a median of 282.

We focus first on the performance of the sampling procedure on an individual tree basis. Four estimators of tree BFV were considered. Estimators $\widehat{Y}_{Ai}$ and $\widehat{Y}_{Bi}$ are alike in that they use *known* merchantable height when partitioning the ith tree into logs, i.e., the actual number of logs in the tree is used to determine the probabilities of selecting logs for measurement. Estimators $\widehat{Y}_{Ci}$ and $\widehat{Y}_{Di}$ are alike in that they use an *estimated* merchantable height derived from a taper function.

Aside from the above difference, estimators $\widehat{Y}_{Ai}$ and $\widehat{Y}_{Ci}$ are alike in that the *known* DIB was used as the scaling diameter of each selected log, which in a field application presumes that the DOB and bark thickness of each selected log can be measured. Similarly, estimators $\widehat{Y}_{Bi}$ and $\widehat{Y}_{Di}$

---

[1] We thank Donald E. Beck, USDA Forest Service, Ashville, NC, for permitting us to use these data in this study.

are alike in that DIB was computed as 0.8900 · DOB where 0.8900 was the average DIB:DOB ratio among the 6636 stem sections.

For a taper function we chose the simple parabolic function used by Gregoire, Valentine, and Furnival (1986):

$$\mathrm{PDOB}(h) = D_B \left(\frac{H_T - h}{H_T - H_B} \right)^{0.5} \qquad [2]$$

where PDOB(h) is predicted outside-bark diameter at height h on the stem, D_B is dbh, and H_B is breast height.

For $\widehat{Y}_{Ci}$ and $\widehat{Y}_{Di}$, it is possible that the real or imputed scaling diameter is less than the minimum scaling diameter. In this situation, we nonetheless scaled the log — a practice which tended to reduce mean square error over the alternative strategy of assigning zero BFV to the log.

The bias (expected minus target value) and root mean square error (RMSE) of each estimator were computed and expressed as a percentage of the target value, Y_i. Expected values were computed on the basis of sampling one log per tree ($m = 1$). These results are summarized in Table 1a. Estimator $\widehat{Y}_A$ is unbiased in these results because the measured DIB was considered as being equivalent to the scaling diameter of the log. The slight negative relative bias of $\widehat{Y}_B$ arises because DIB computed by adjusting DOB occasionally is less than scaling diameter. This phenomenon is a result jointly of the taper function that we used and the manner of computing DIB from DOB. For these data, no pattern was evident: butt logs did not seem any more or less prone to underscaling than intermediate and top logs.

The positive bias of $\widehat{Y}_C$ and $\widehat{Y}_D$ reflects the tendency for the taper function, [2], to exaggerate merchantable height for these trees. The average relative bias is only a small portion of the average relative RMSE. The increased variability of $\widehat{Y}_C$ and $\widehat{Y}_D$ over $\widehat{Y}_A$ and $\widehat{Y}_B$ is due mainly to the imprecision in determining the number of logs in the tree from a taper function.

For each estimator we graphed the 223 values of relative bias against tree BFV, in order to determine if an estimator's performance varied with tree size. Similar graphs were prepared with relative RMSE as the ordinate. Relative bias decreases in absolute magnitude with increasing tree BFV for $\widehat{Y}_B$, $\widehat{Y}_C$, and $\widehat{Y}_D$. For all four estimators, relative RMSE tends to increase with increasing BFV (Figure 1). The reason for this pattern is unclear.

We also examined the performance of the sampling procedure and various estimators using the DOB taper function developed from these same yellow poplar data by Knoebel, Burkhart, and Beck (1984):

$$\mathrm{PDOB}(h) = 1.24452 \cdot D_B^{0.94785} H_T^{-0.77189} (H_T - h)^{0.76900} \qquad [3]$$

The sampling results are shown in Table 1b. The better fit provided by this taper function to the actual tree profiles is reflected in the sizeable reductions in average root mean square error for estimators $\widehat{Y}_C$ and $\widehat{Y}_D$. These reductions are due mainly to the ability with [3] to predict the correct number of logs in the merchantable portion of the stem more often than with [2]. The average bias and RMSE using [3] did not change much for estimators $\widehat{Y}_A$ and $\widehat{Y}_B$ because the taper function with these estimators is used only to determine the approximate scaling diameters for the correctly known number and length logs.

The RMSE values in Table 1 compare favorably to the summary statistics of model prediction error reported by Burk, Hans, and Wharton (1989) for a board-foot volume equation fitted to 180 soft maple and yellow poplar trees together. Analogous to our estimators $\widehat{Y}_{Ai}$ or $\widehat{Y}_{Bi}$ is their fitted model which uses actual tree height and for which they reported an average absolute relative error of 14%; analogous to our estimators $\widehat{Y}_{Ci}$ or $\widehat{Y}_{Di}$ is their fitted model which uses an ocularly estimated height, for which they reported an average absolute relative error of 25.3%. These errors were obtained on 180 trees fitted to a nonlinear form-class volume equation.

TABLE 1. Average relative bias and root mean square error (RMSE) of four estimators of standing tree board-foot volume. In a), sampling was conducted with the parabolic taper function of Gregoire, Valentine, and Furnival (1986). In b), sampling was conducted with the taper function of Knoebel, Burkhart, and Beck (1984).

		$\widehat{Y}_{Ai}$	$\widehat{Y}_{Bi}$	$\widehat{Y}_{Ci}$	$\widehat{Y}_{Di}$
a)	Bias (%)	0.0	-4.0	5.3	2.3
	RMSE (%)	8.6	11.1	20.9	20.9
b)	Bias (%)	0.0	-4.0	-2.0	-5.1
	RMSE (%)	8.6	11.1	12.8	14.6

We next examined the performance of a two-stage sampling in which trees are selected in the first stage with probability proportional to size with replacement. The measure of size was $D_B^2 H_M$ or $D_B^2 \widehat{H}_M$, where H_M is known merchantable height and $\widehat{H}_M$ is merchantable height predicted by the taper function used in the second stage of sampling. The former measure of size was used in conjunction with second-stage estimates $\widehat{Y}_{Ai}$ and $\widehat{Y}_{Bi}$ that use known merchantable height in the partition of the tree into logs, whereas the latter measure was used in conjunction with $\widehat{Y}_{Ci}$ and $\widehat{Y}_{Di}$. The bias and mean square error of the two-stage estimator, for a single tree sample at stage one ($n = 1$) and a single log sample at stage two ($m = 1$), was evaluated and expressed relative to the total BFV of the 223 trees (Table 2).

These results, too, compare favorably with the results of Burk, Hans, and Wharton that we quoted above, despite the fact that those in Table 2 are based on a single observation. For a first-stage sample of $n = 25$ trees with a single second-stage sample per tree, RMSE values in Table 2 reduce to 3.5% and 5.1% for $\widehat{Y}_A$ and $\widehat{Y}_B$, respectively, and to 6.4% and 5.5% for $\widehat{Y}_C$ and $\widehat{Y}_D$ with the simple taper function, and to 4.9% and 6.5% for the taper fuction tailored to these yellow poplar data.

For this small sample, bias is a small proportion of total error. It is interesting to note that despite its bias, estimator $\widehat{Y}_B$ has smaller RMSE than estimator $\widehat{Y}_A$ which uses measured DIB in the scaling step of the sampling procedure. No reason for this result is readily discernible. We note, too, the sizeable reduction in RMSE achieved with the better fitting taper function of Knoebel, Burkhart, and Beck.

TABLE 2. Relative bias and RMSE of two-stage sampling estimators of total board-foot volume evaluated for a first-stage sample size $n = 1$ tree and second-stage sample size $m = 1$ log. In a), sampling was conducted with the parabolic taper function of Gregoire, Valentine, and Furnival (1986). In b), sampling was conducted with the taper function of Knoebel, Burkhart, and Beck (1984).

		$\widehat{Y}_A$	$\widehat{Y}_B$	$\widehat{Y}_C$	$\widehat{Y}_D$
a)	Bias (%)	0.0	4.0	2.6	-1.0
	RMSE (%)	17.6	16.0	29.6	27.0
b)	Bias (%)	0.0	4.0	1.5	5.2
	RMSE (%)	17.6	16.0	21.7	20.1

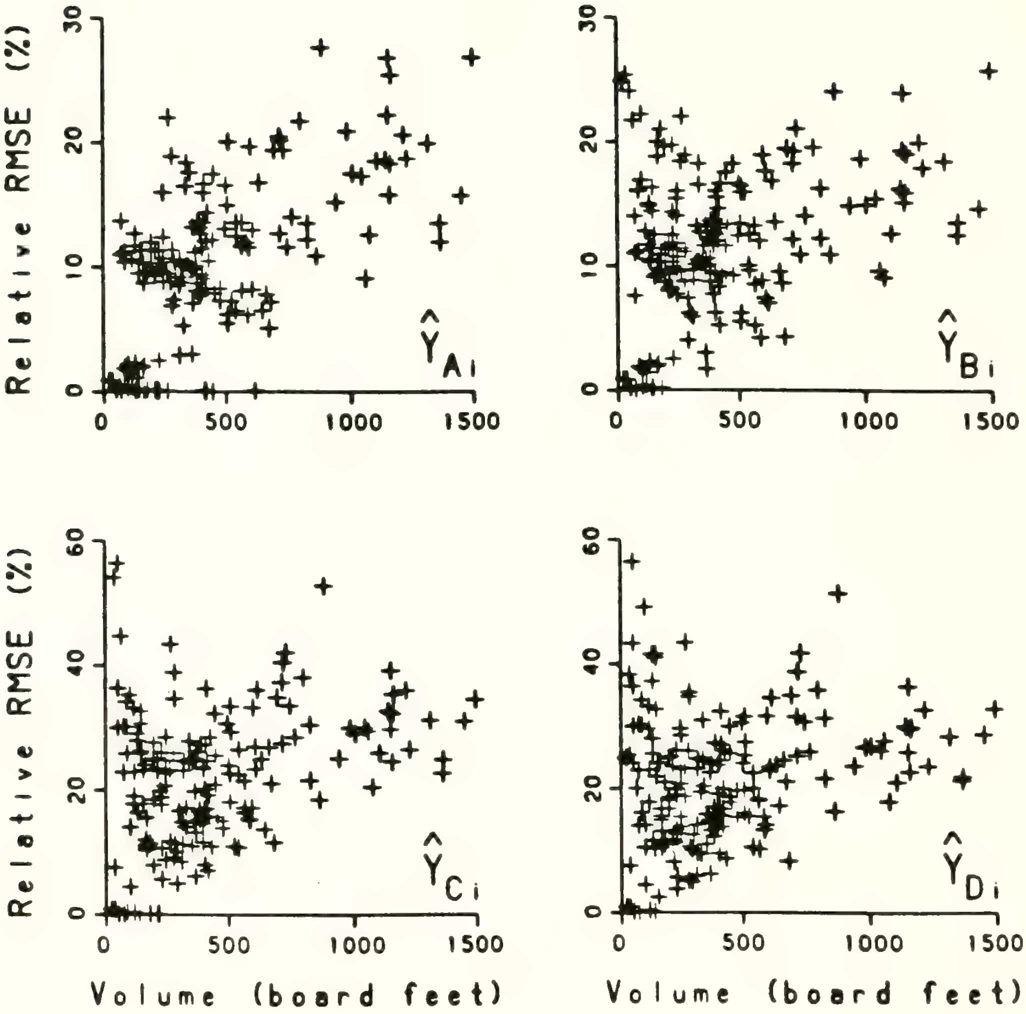


Figure 1. Relative RMSE of four estimators of standing tree board-foot volume. Sampling was conducted with the taper function of Gregoire, Valentine, and Furnival (1986).

DISCUSSION

In our previous work to develop sampling procedures to estimate tree biomass (Valentine, Tritton, and Furnival 1984), stem cubic-foot volume and increment (Gregoire, Valentine, and Furnival 1986, 1987), and log cubic-foot volume (Furnival, Valentine, and Gregoire 1986), we have used a taper function in a different fashion than we have proposed here. The common thread in all these procedures, however, is the use of a function to determine sample selection probability density or mass. In the present case, logs within the tree can be selected with a uniform probability, thereby dispensing with the need for a taper function. Such a practice is not recommended because of the increased efficiency which accompanies the use of a taper function — even one of such general applicability as [2] — to determine selection probabilities proportional to log volume. PPS sampling has an important practical advantage, too, of increasing the likelihood of scaling the more accessible butt log of the tree, rather than an upper-stem log. For the yellow-poplar trees, the average probability of the butt log was 0.49 for both taper functions we examined. Of course, upper-stem logs will be selected on occasion.

To sample for BFV, we suggest that a taper function may be used, too, to estimate the merchantable height of a tree in order to partition the tree stem into logs. The precision of the resulting estimator depends greatly on the precision with which the number of logs can be determined.

Despite its use of a taper function in the manner described above, the estimator of tree BFV makes no assumption about the actual taper of the tree stem, nor that all trees in a stand have a common form. Furthermore, extensive dendrometry of the stem is not required to measure taper or taper-steps.

Grosenbaugh (1964) suggested that local sampling could help to establish appropriate constants to relate primary measurements of V-S-L directly to lumber yield, thereby sidestepping the use of scaled volumes altogether. Our approach differs because it aims at the BFV of a tree that would be scaled on its logs, bucked and on the ground.

REFERENCES

Avery, T. E. and Burkhart, H. E., 1983: Forest Measurements. 3rd ed. McGraw-Hill Book Co., New York, 331 pp.

Beck, D. E., 1963: Cubic-foot volume tables for yellow-poplar in the southern Appalachians. U.S.D.A. Forest Service Research Note SE-16, Asheville, NC, 4 pp.

Burk, T. E., Hans, R. P., and Wharton, E. H., 1989: Individual tree volume equations for the northeastern United States: evaluation and new form quotient board foot equations. Northern Journal of Applied Forestry, vol. 6, no. 1, pp. 27-31.

Burkhart, H. E., Farrar, K. D., Amateis, R. L., and Daniels, R. F., 1987: Simulation of individual tree growth and stand development in loblolly pine plantations on cutover, site-prepared areas. Publication No. FWS-1-87, School of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University, Blacksburg, VA, 47 pp.

Furnival, G. M., Valentine, H. T., Gregoire, T. G., 1986: Unbiased estimation of log volume by importance sampling. Forest Science, vol. 32, no. 4, pp. 1073-1078.

Gregoire, T. G., Valentine, H. T., and Furnival, G. M., 1986: Estimation of bole volume by importance sampling. Canadian Journal of Forest Research, vol. 16, pp. 554-557.

Gregoire, T. G., Valentine, H. T., and Furnival, G. M., 1987: Sampling methods for estimating stem volume and volume increment. Forest Ecology and Management, vol. 21, pp. 311-323.

Grosenbaugh, L. R., 1964: Some suggestions for better sample-tree-measurement. Proceedings Society American Foresters meeting, October 20-23, 1963, Boston, Massachusetts, pp. 36-42.

Hann, D. W. and Bare, B. B., 1978: Comprehensive tree volume equations for major species of New Mexico and Arizona. U.S.D.A. Forest Service Research Paper INT-209, Ogden, Utah, 43 pp.

Hilt, D. E., Rast, E. D., and Bailey, H. J., 1983: Predicting diameters inside bark for 10 important hardwood species. U.S.D.A. Forest Service Research Paper NE-531, Broomall, PA, 7 pp.

Knoebel, B. R., Buckhart, H. E., and Beck, D. E., 1984: Stem volume and taper functions for yellow-poplar in the southern Appalachians. Southern Journal of Applied Forestry, vol. 8, no. 4, pp. 185-188.

Mawson, J. C. and Connors, M. E., 1987: Girard form class equations. Northern Journal of Applied Forestry, vol. 4, no. 1, pp. 27-31.

Saucier, J. R., Phillips, D. R., and Williams, Jr., J. G., 1981: Green weight, volume, board-foot, and cord tables for the major southern pine species. Georgia Forest Research Paper No. 19, Georgia Forestry Commission, Macon, GA, 63 pp.

Valentine, H. T., Tritton, L. M., and Furnival, G. M., 1984: Subsampling trees for biomass, volume, or mineral content. Forest Science, vol. 30, no. 3, pp. 673-681.

Estimating the Volume of a Radiata Pine Stand Using Importance Sampling

Harry V. Wiant, Jr.
Visiting Fellow,
Department of Forestry, Australian National University,
Canberra, ACT 2601, Australia.
(Professor of Forestry, West Virginia University, Morgantown, WV 26506, USA.)

and

Geoffrey B. Wood,
Senior Lecturer,
Department of Forestry, Australian National University,
Canberra, ACT 2601, Australia.

ABSTRACT

A two-phase sampling scheme was used to estimate the merchantable volume (to 10 cm top dub) of a sample plot of 36-year-old, plantation-grown radiata pine (*Pinus radiata* D. Don) comprising 55 trees. The first phase consisted of visiting each tree in the plot, estimating its volume by eye, and making a 3P selection of sample trees. The second phase involved estimating the volume of each 3P sample tree using a proxy taper function and adjusting each volume estimate by the ratio of measured to estimated cross-sectional area at a point on the tree bole, the height of which was determined by importance sampling. The estimate of total merchantable volume derived using importance sampling was comparable to that obtained with detailed dendrometry of the 3P sample trees while reducing the dendrometry by about 96%. When applied to trees for which volume functions are not available, importance sampling combined with 3P sampling, or any other type of sampling procedure, has great potential for minimising the cost of forest inventory.

INTRODUCTION

Grosenbaugh (1967, 1974) introduced 3P sampling for forest inventory by which rapid volume estimates are made for individual trees and a few are selected for detailed measurement (e.g., by dendrometry) with a probability of selection proportional to predicted volume. The principal advantages of this sampling method are its efficiency and the elimination of bias arising from volume tables.

Gregoire *et al.* (1986) suggested that importance sampling could be a practical alternative to detailed dendrometry in inventories involving 3P sampling. Importance sampling requires use of a proxy taper function to estimate the volume of the tree bole. This is then adjusted by the ratio of actual cross-sectional area 'A' at a randomly selected point 'h' on the bole to the cross-sectional area 'a' at that point estimated from the proxy taper function. The height h is dependent upon the distribution of estimated volume along the bole, a procedure called "importance sampling".

As a simple example, consider an artificial tree of conoid shape of basal radius 0.1 m and total height 10.0 m. The actual volume V_a of this tree is 0.105 m^3 and the actual taper function is : $r = 0.1 - 0.01h$, where r and h are the radius (m) and height above ground (m) respectively of any point on the bole. In the real world, we would know neither V_a nor the actual taper function. However, we can use a proxy taper function, viz.: $r = 10 - h$ .... (Eqn. 1)

This is obviously a rough proxy function but it will serve well for this simple illustration. In practice, the proxy taper function need not be the most appropriate to preserve unbiasedness, but precision is improved the better the approximation (Gregoire *et al.*, 1986). Using the proxy function (Eqn. 1), the estimated volume (V_e) of the tree by integration is:

$$\begin{aligned} V_e &= \pi\int_0^{10} r^2 \partial h \\ &= \pi\,[(100(10) - 10(10)^2 + 1/3(10)^3) - (100(0) - 10(0)^2 + 1/3(0)^3)] \quad \text{.................. (Eqn. 2)} \\ &= 1047\ \text{m}^3 \end{aligned}$$

To select h, the point where a cross-sectional area measurement is needed such that the probability is 0.5 it will occur at a point below which 50% of the estimated volume occurs, 0.6 at a point below which 60% of the volume occurs, etc., requires the prior selection of a uniform random number u [0, 1]. Suppose the random number we generate is 0.295. In this case, we must find the h-value below which the volume of the bole is 0.295 x 1047 = 309 m^3. Thus, we can write : $\pi\,[100h - 10h^2 + 1/3h^3] = 309\ \text{m}^3$. A simple programmable pocket calculator can be used to solve for h by iteration (refer to program in the Appendix). In this instance, a solution of 1.1 m is obtained.

As indicated earlier, the true taper function for the tree is: $r = 0.1 - 0.01\ h$. Thus, the value we would measure for r, at h = 1.1 m, is : $r = 0.1 - 0.01(1.1) = 0.089$ m, and the cross sectional area (A) at h is : $\pi\,(0.089)^2 = 0.0249\ \text{m}^2$.

From the proxy taper function (Eqn. 1), $r = 10 - h$, i.e., 10 - 1.1 = 8.9 m, for which the estimated cross sectional area (a) is : 248.846 m^2. The importance volume estimate (V_i) is then : $V_i = V_e\ (A/a) = 1047\ (0.0249/248.846) = 0.105\ \text{m}^3$. With our artificial tree, not plagued by the variability in shape along the tree bole found in nature, the importance estimate of volume is exact.

The aim of the study described in this paper was to compare estimates of stand volume derived using pure 3P sampling with, in one case, conventional detailed dendrometry of sample trees and, in the other, importance sampling which requires minimum dendrometry of these trees.

SELECTING THE PROXY TAPER FUNCTION

Proxy taper functions selected for testing included two functions developed for radiata pine (*Pinus radiata* D. Don) in New Zealand by Goulding and Murray (1975) and Lowell (1986), and the "generic" function of Gregoire *et al.* (1986) which makes the reasonable assumption that the square of the radius decreases linearly with h (i.e., bole shape is a 2nd-degree paraboloid). For comparative purposes, an obviously poor proxy function (Eqn. 1), which assumes that the radius of the bole decreases linearly with height above ground (i.e., solid shape is a cone), was also selected.

These four taper functions were tested using data drawn from a radiata pine sample tree data base held in the Department of Forestry, ANU. The particular data set selected comprised 114 trees ranging in age from 21 to 45 years, dbhob from 17 to 56 cm, and total height from 13 to 34 m. Data on each tree included dbhob, total height, merchantable height to a 10 cm top diameter under bark (dub), diameters under bark at close intervals from stump height to breast height (1.3 m) and at about 1 m intervals (or closer) between breast height and the tip, and total and merchantable bole volume under bark (to 10 cm dub) both derived using Smalian's formula. An average of 23.4 separate volume estimates using importance sampling was made per tree. For each h-value, the volume estimate derived from a particular taper function was adjusted by the A/a-ratio. This gave a total of 2692 estimates for analysis.

The results presented in Table 1 indicate that the mean square error of the volume estimate was least but the relation between bias and relative height on the bole was greatest for the equation developed by Gregoire *et al.* (1986). Overall, the equation of Goulding and Murray (Table 2) performed best. In particular, it had the lowest mean square error for the ratios of measured volume to estimated volume which is so important in 3P sampling. For this reason, it was selected for use in the field trial.

THE FIELD TRIAL

The forest stand selected for the trial was located in Cpt. 137 Stromlo Forest, A.C.T. It comprised a rectangular sample plot of 55 trees of 36-year-old radiata pine. The plot and surrounding stand had recently received a second commercial thinning.

Based on previous experience in similar exercises, we assumed the coefficient of variation (c) of the ratios of measured to estimated volume of the 3P sample trees would be 16%. Thus, the size (n) of the 3P sample required to give a sampling error (e) of ± 5% 2 times in 3 (Student's t = approx. 1) was :

$$n = (c\ t\ /\ e)^2 = ((16 \times 1)\ /\ 5)^2 = 10.2$$

A random number list for the stand was then generated using the procedure described by Wood (1988) after which pure 3P sampling of the 55 trees was applied. This required an ocular estimate of the merchantable volume under bark of each tree and selection of 3P sample trees based on the random number list. A total of 11 3P-trees was selected. On each of these trees, diameter over bark and bark thickness at breast height (1.3 m) were measured by tape and bark gauge respectively, height to a 10 cm top was estimated by Spiegel Relaskop, and measurements of diameter over bark and height at 8 to 10 points along the bole including the tip were made with a Barr and Stroud FP-15 rangefinder dendrometer. Diameter over bark was measured by tape at one importance height (h) per tree if accessible from the ground (7, 3P-trees), otherwise by both Relaskop and Barr and Stroud dendrometer (4, 3P-trees). The simple program used to generate h is given in the Appendix.

The 3P processing package STXMOD (Chehock 1982) was used to process the 3P data and derive estimates of stand volume underbark using bark option-2 of Grosenbaugh (1967). The resulting volume estimates are expected to be negatively biassed by 7.6 % due to known bias of bark option-2 for radiata pine in the A.C.T. (Johnson and Wood 1987), but this does not invalidate a comparison of the stand volume estimates derived using detailed dendrometry and importance sampling. However, to permit a wider comparison, parameters of the 55 trees in the stand were also measured to permit total merchantable volume underbark to be derived using both the 4-variable regional volume table for *Pinus radiata* in South Australia (Lewis *et al.*, 1973) and the 2-variable volume table for *Pinus radiata* in Stromlo Forest, A.C.T. (Bary and Borough 1980).

RESULTS

Estimates of total merchantable volume underbark (to 10 cm top dub) of the stand derived from importance sampling were comparable with those derived from detailed dendrometry using the Barr and Stroud dendrometer, irrespective of whether a Barr and Stroud or a Spiegel Relaskop was used to derive the diameter over bark at the importance sampling point. However, the sampling errors were slightly larger for importance sampling (Table 3). When adjusted for the known bias in Grosenbaugh's bark function (option-2), the volumes derived from detailed dendrometry and importance sampling are consistent with those

Table 1. Statistics comparing volume estimates derived from simulated importance sampling (N = 2692).

Statistic	True volume (measured directly)	Volume derived from importance sampling			
		Goulding & Murray	Lowell	Gregoire et al.	Conoid
IMPORTANCE VOLUME					
Mean (cu. m)	0.443	0.446	0.455	0.437	0.519
Std. Dev. (cu. m)	0.361	0.377	0.396	0.354	0.628
Bias (cu. m)	---	0.004	0.012	-0.005	0.076
Mean sq. error	---	0.142	0.157	0.125	0.400
Student's-t (Ho : Bias = 0)	---	2.861**	6.858***	-3.086**	10.126***
RELATIONSHIP BETWEEN BIAS AND RELATIVE HEIGHT					
r-squared	---	0.11***	0.31***	0.32***	0.27***
RATIO OF MEASURED TO ESTIMATED VOLUME					
Mean	---	1.003	0.993	1.021	0.966
Std. Dev.	---	0.083	0.109	0.137	0.238
Mean sq. error	---	0.007	0.012	0.019	0.058
Coeff. of Var.(%)	---	8.3	11.0	13.4	24.6
Student's-t (Ho : Ratio = 1)	---	2.006*	-3.366**	7.908***	-7.413***

* = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$

Table 2. Equations of Goulding and Murray (1975)

Variable	Relevant equation
Diameter	$d^2 = (C - E + F - G + I)\,P$
Volume	$V = \left[Q\,(A - B + J - K + M)\right]_{LL}^{LU}$
Height on bole	$V(LL, h) = \mu V(LL, LU)$ (Solve for h by iteration)

where:

d = diameter inside bark (cm)

$C = 22.686\,(L/H)^5$

$E = 44.310\,(L/H)^4$

$F = 26.708\,(L/H)^3$

$G = 3.5452\,(L/H)^2$

$I = 1.1714\,(L/H)$

L = distance from tip of the tree (m)

H = total tree height (m)

$P = 0.33021\,D^2 + 17.070$

D = dbh outside bark (cm)

V = volume inside bark (m^3)

$Q = 0.0000259346\,D^2 + 0.00134067$

$A = 3.781\,H^{-5}\,L^6$

$B = 8.862\,H^{-4}\,L^5$

$J = 6.677\,H^{-3}\,L^4$

$K = 1.182\,H^{-2}\,L^3$

$M = 0.586\,H^{-1}\,L^2$

LU = distance from tree tip to stump (m)

LL = distance from tree tip to merchantable top (m)

μ = random number from a uniform [0,1] distribution

h = randomly selected height for cross-sectional area measurement

Table 3. Statistics comparing estimates of total stand volume derived from detailed dendrometry and importance sampling of 3P sample trees.

Statistic	Detailed dendrometry (Barr & Stroud)	Importance sampling[1]	
		Barr & Stroud	Relaskop
Total volume (cu. m)	47.4	48.6	49.3
Std. error (%)	5.1	5.3	5.8
Coeff. Var. (%)	18.8	19.6	21.5
95% conf. limits (cu. m)	42.0 - 52.7	43.0 - 54.4	42.9 - 55.7

[1] Using taper equation of Goulding and Murray (1975).

derived from the two volume tables, viz.:

3P estimate (detailed dendrometry)	51.3 m^3 ± 5.1 %
Importance sample estimate (Barr & Stroud)	52.6 m^3 ± 5.3 %
Importance sample estimate (Relaskop)	53.4 m^3 ± 5.8 %
South Australia *P. radiata* volume table (4-variable)	52.1 m^3
Stromlo (A.C.T.) *P. radiata* volume table (2-variable)	52.6 m^3

The Spiegel Relaskop appeared to be adequate for taking a single upper stem measurement of diameter as required in importance sampling but this needs to be tested in a larger sample.

Importance sampling in this trial reduced the upper-stem measurements required to be made by dendrometry from 108 in the conventional 3P sample (which requires detailed dendrometry) to 4 only, i.e., by about 96%, with only a slight increase in sampling error (Table 3).

The power of the importance adjustment is well demonstrated by results obtained assuming a conoid taper function for the tree boles (refer to the earlier illustration in the Introduction). Without the importance adjustment, the volume estimate would have been 654 354 m^3, a huge overestimate; with the adjustment, it was 50.5 m^3, which is a reasonable estimate (cf. results in Table 3).

DISCUSSION

Importance sampling should prove useful for estimating the volume of forests containing species for which acceptable taper functions are not available. Should assortment of the total volume estimate into log grades or wood products be required, it can be done by measuring on each sample tree the height to each point where the grade or product changes. Volumes by grade or product within a given tree can then be determined by integration and adjusted subsequently by the A/a-ratio. The program in the Appendix can be used to do this by calling the height to the top of a given grade or product section "MERCH HT" and height to the bottom of the section "STUMP HT". Similarly, should merchantable volume estimates to different top diameters under bark be required, one would need to estimate (e.g. by Relaskop) the heights to the various merchantable limits and proceed as outlined above.

The considerable reduction in the dendrometry required when applying importance sampling is due to two factors. Firstly, only one measurement on the upper-stem is required per sample tree (although more can be taken - refer Gregoire *et al.* (1986)) and, secondly, based on our experience in this trial, more than half of the measurements are within easy reach from the ground, i.e., they can be measured directly, particularly if a light 1-2 m aluminium ladder is carried in the field.

The importance sample point h in some cases will fall on an unrepresentative point, e.g., on a whorl, or the bole of the tree at h will not be visible because of obscuring branches or foliage. In each case, we recommend that the point of measurement be changed to the nearest representative, visible point on the tree bole. Should the tree fork below the importance sample point, either measure by dendrometry the diameter at h on each leader and derive the quadratic mean diameter, or calculate new 'h's until one is found which falls below the breach.

We wish to emphasise two points:

1. A great deal of time and effort should not be spent finding or establishing the "best" taper function for trees in a particular forest. The function presented by Gregoire *et al.* (1986) assumes a solid shape for the tree bole of second degree paraboloid. As the great majority of the main boles of trees around the world closely approximate this shape, Gregoire *et al.*'s function should produce acceptable results for any species in any location.

2. Importance sampling is not confined to use with 3P sampling. It can be used in any sampling design where sample trees are taken, particularly trees of species for which reliable volume functions are not available. It has great potential for reducing the cost of forest inventory especially inventories using variable probability sampling procedures, e.g., list, point-3P, point-model based sampling, etc. We feel that advances in sampling theory and technology are rapidly approaching the stage where reduction of field measurements to a bare minimum is achievable while still maintaining the reliability of the estimates derived from inventory.

ACKNOWLEDGEMENTS

We thank Dr T.G. Gregoire for helpful comments, Dr L.T. Carron for reviewing the manuscript, Mr J.A. Miles for computational assistance and Mrs Jeanne Wiant for assistance.both in the office and field.

REFERENCES

Bary, G.A.V. and Borough, C.J., 1980: Tree volume tables for *Pinus radiata* in the Australian Capital Territory. CSIRO Division of Forest Research Report No. 11, 19 pp. (Unpublished).

Chehock, C.R., 1982: Timber volume determination with multi-stage sampling: the application of 3-P sampling and the STXMOD program. Technical Publication SA-TP 19, USDA Forest Service, Atlanta, GA 30367. 135 pp.

Goulding, C.J. and Murray, J.C., 1975: Polynomial taper equations that are compatible with tree volume equations. New Zealand Journal of Forest Science, vol. 5, pp. 313-322.

Gregoire, T.G., Valentine, H.T. and Furnival, G.M., 1986: Estimation of bole volume by importance sampling. Canadian Journal of Forest Research, vol. 16, pp. 554-557.

Grosenbaugh, L.R., 1967: STX-FORTRAN-4 PROGRAM for estimates of tree populations from 3P sample-tree-measurements. USDA Forest Service Research Paper PSW-13, 76 pp.

Grosenbaugh, L.R., 1974: STX 3-3-73: tree content and value estimation using various sample designs, dendrometry methods, and V-S-L conversion coefficients. USDA Forest Service Research Paper SE-117, 16 pp.

Johnson, T.S. and Wood, G.B., 1987: Simple linear model reliably predicts bark thickness of radiata pine in the Australian Capital Territory. Forest Ecology and Management, vol. 22, pp. 173-183.

Lewis, N.B., Mc Intyre, G.A. and Leech, J.W., 1973: Regional volume table for *Pinus radiata* in South Australia (metric edition). Woods and Forests Dept., South Australia. 64 pp.

Lowell, K.E., 1986: A flexible polynomial taper equation and its suitability for estimating stem profiles and volumes of fertilized and unfertilized radiata pine trees. Australian Forest Research, vol. 16, pp. 165-174.

Wood, G.B., 1988: Generating the list of random numbers for 3P sampling. Australian Forestry, vol. 50, pp. 260-264.

APPENDIX

Program in BASIC for finding the importance height and calculating the importance volume for sample trees of radiata pine. The program is written for a Sharp Scientific Pocket Calculator PC-1403 or EL-5500 III and incorporates the bark function of Johnson & Wood (1987) and the taper function of Goulding & Murray (1975).

PROGRAM		COMMENTS
10	INPUT "1 FOR HT, 2 FOR VOL "; X	
20	L=0	
30	INPUT "DBH "; D	
40	INPUT "TOT HT "; H	
50	INPUT "MERCH HT "; MH	
60	INPUT "STUMP HT "; SH	
70	IF X=1 THEN 90	
80	IF X=2 THEN 110	
90	INPUT "RAN NO "; U	Input a random number u (0,1)
100	GOTO 290	
110	INPUT "2*BARK "; BK	Double bark thickness at breast height
120	INPUT "IMP HT "; HT	Height at which the importance diameter is measured
130	INPUT "IMP D "; ID	Diameter over bark at the importance height
140	XI=3.141592654	Value for π
150	RD=ID/D	) Estimates bark thickness at the importance height using the
160	BT=(0.12265+0.84116*RD^4)*BK/2	) equation of Johnson & Wood (1987)
170	DB=ID-2*BT	Corresponding dub at the importance height
180	A=(XI/40000)*DB^2	Sectional area ub at the importance height
190	L=H-HT	
200	C=22.686*(L/H)^5	)
210	E=44.310*(L/H)^4	) Coefficients for estimating dub from the proxy taper function of
220	F=26.708*(L/H)^3	) Goulding & Murray (1975)
230	G=3.5452*(L/H)^2	)
240	I=1.1714*(L/H)	)
250	P=0.33021*D^2+17.07	)
260	ED=((C-E+F-G+I)*P)^0.5	Dub at the importance height estimated from the proxy taper function
270	IA=(XI/40000)*ED^2	Corresponding sectional area under bark
280	RAT=A/IA	Ratio of sectional areas
290	LL=H-MH	
300	LU=H-SH	
310	L=LL	
320	GOSUB 530	
330	VL=VOL	
340	L=LU	
350	GOSUB 530	
360	VU=VOL	
370	V=VU-VL	Integrated volume derived from the proxy taper function
380	IF X=1 THEN 420	
390	VI=V*RAT	Corrects the integrated volume
400	PRINT "IMP VOL= "; VI	Prints the importance volume
410	GOTO 10	
420	VN=U*V	
430	L=H	
440	L=L-0.1	
450	GOSUB 530	
460	V=VU-VOL	
470	IF ABS(VN-V)<=0.005 THEN 490	Test to terminate the iteration
480	GOTO 440	
490	HT=H-L	
500	PRINT "HT= "; HT	Prints the importance height
510	GOTO 10	
520	END	
530	Q=0.0000259346*D^2+0.00134067	)
540	A=3.781 * H^-5 * L^6	)
550	B=8.862 * H^-4 * L^5	) Coefficients of the proxy taper function for estimating volume
560	J=6.677 * H^-3 * L^4	)
570	K=1.182 * H^-2 * L^3	)
580	M=0.586 * H^-1 * L^2	)
590	VOL=Q*(A-B+J-K+M)	Derives the volume from the proxy taper function
600	RETURN	

Biomass Estimates for Tropical Forests Based on Existing Inventory Data

Andrew J. R. Gillespie
Biometrician
USDA Forest Service, Forest Pest Management
Fort Collins, CO 80525

Sandra Brown
Associate Professor of Forest Ecology
University of Illinois at Champaign-Urbana
Urbana, IL 61801

Ariel E. Lugo
Project Leader
USDA Forest Service Institute of Tropical Forestry
Rio Piedras, PR 00928-2500

ABSTRACT

Biomass estimation is becoming important for a variety of planning and scientific applications in many tropical countries. Biomass inventories in many areas of the world are rare or nonexistent, so that biomass estimates must be based on other types of information. Most countries have some inventory-based estimates of forest volumes. We have developed a toolbox of approaches for making point and interval estimates of total aboveground biomass based on various types of forest volume inventories. The toolbox includes (1) methods of converting commercial volume to total aboveground biomass, (2) regression equations to estimate biomass directly from stand tables, and (3) methods of estimating the numbers of stems or total biomass in small, missing diameter classes based on known values for larger diameter classes. While not as reliable as a well planned biomass inventory, our estimates are expected to be more reliable than current estimates based on direct measurement of relatively few, non-randomly selected plots.

INTRODUCTION

Biomass inventories are becoming increasingly relevant for tropical forest resource assessment. Weight is a more direct measure of quantity than volume for many applications, for example inventories of fuelwood and pulp. Biomass is also more relevant in scientific studies focusing on such things as forest productivity and nutrient cycling. Our research into tropical biomass estimation methods has been part of a larger effort aimed at quantifying the effect of land use change in the tropics on the global carbon cycle.

While most of the increase in atmospheric carbon is directly attributable to fossil fuel combustion, the global carbon budget remains unbalance. In 1980, the current reference year for balancing the global carbon budget, the documented sources of carbon (fossil fuel combustion and cement production) are 5.3 billion metric tons (BMT) per year, while the atmospheric and ocean sinks are estimated to absorb no more than 4.8 BMT per year. Oceanographers hypothesized that the remaining carbon was taken up by terrestrial ecosystems (Trabalka, 1985). Ecologists argued this was unlikely, and it was more likely that conversion of tropical forests into lower-biomass secondary forests and pastures acted as an additional, unaccounted-for source of carbon (Houghton et al., 1983). A project is currently underway with the goal of modeling the carbon flux associated with this changing tropical land use. These models depend on two critical estimates: the rates of land use change, i.e. numbers of hectares per year being changed from one land category to another, and the total biomass per unit area of the various land categories. Given these two factors, it is a simple matter to multiply the numbers of hectares times the biomass per hectare and estimate the net loss or gain of carbon due to land use change in a given geographical region. We focus on the total aboveground biomass of forested land because forests account for most of the carbon, and because these estimates are the most uncertain. The purpose of our research is to determine point and interval estimates of biomass for various types of disturbed and undisturbed tropical forests.

The ideal approach to this problem would be to design a worldwide biomass inventory, with permanent plots, whole-tree sampling, and stratification by forest type and land use. This is impossibly expensive for a single research project. The United Nations Food and Agriculture Organization (FAO) is currently conducting their second tropical forest assessment, but they too face limitations of time and funding, so that they rely heavily on in-country inventories, careful mapping, interviews of information gatherers, and as much travelling as possible (K. D. Singh, 1989, personal communication).

The earlier models of carbon flux from tropical forests used biomass estimates based upon ecological studies (e.g. Ovington and Olson (1970) and Hozumi et al. (1969)). These data were based on destructive sampling of relatively few plots in various types of tropical forests. The purpose of these studies was to focus on ecosystems: the stocks and flows of energy, nutrients, and mass within a single location. The purpose was not to estimate biomass for a given forest, but rather to use biomass as a measure of ecosystem flows. The studies generally focused intensively on a single location in great depth, rather than extensively over an entire forest. This single location was generally not randomly selected, but was purposefully chosen to be relatively diverse and vigorous so as to best illustrate the phenomena of interest. Although satisfactory for their ecosystem studies, these data are not appropriate for estimating the mean biomass/ha of an entire country or continent. These estimates generally overestimate the mean biomass/ha over entire forests (Brown et al., 1989).

We attempt to synthesize more general estimates based on existing forest inventory, using as statistically rigorous a manner as possible, documenting our assumptions as necessary and most importantly attempting to report some measure of error along with our point estimates. This error, expressed as a confidence interval, is extremely important for purposes of analyzing the sensitivity of carbon flux models. Since the available inventory data do not generally address biomass, and the data collection and reporting methods are extremely diverse, no single method is going to work in all cases. We seek to develop a variety of techniques for generating point and interval estimates of biomass based on many types of auxiliary information.

Others working in temperate forests have encountered this problem. Since (i) most of the inventories include information regarding volume, and (ii) volume and biomass are strongly related, one solution is to derive biomass estimates from volume estimates. Johnson and Sharp (1983) did this for temperate forests by estimating the ratio of total aboveground biomass to commercial volume for a wide variety of forests types. Brown and Lugo (1984) used a similar approach for tropical forests, computing a mean-of-ratios estimator without addressing the bias or error. Their estimates, substantially lower than earlier estimates based on ecological data (Brown and Lugo 1982), drew attention to the uncertainty in existing biomass estimates as used in modeling the global carbon flux.

BIOMASS ESTIMATION METHODOLOGY

There are two basic types of data available from tropical forest inventories: stand or stock tables (numbers of trees or volume by diameter class), and unit area estimates, i.e. sums of volume, basal area, or numbers of trees per hectare for a given forested area. We use a regression-based method for stand tables, and a ratio estimator approach for unit area volume data.

Biomass Regressions for Use With Stand Tables

The tropics is diverse region comprised of many different forest types. Given the scarcity of data available to fit biomass regressions for the tropics, it was necessary to make some gross groupings. Brown and Lugo (1982) reported that volume and biomass seemed related to various environmental variables, specifically precipitation and temperature. Using the Holdridge life zone system of ecological habitat classification (Holdridge 1967) and the ratio of temperature to precipitation, they divided tropical forests into three broad groupings: wet ($t/p < 0.007$), moist ($0.007 < t/p < 0.017$), and dry ($t/p > 0.017$). We use a similar breakdown and developed a set of equations for each of these three broad ecosystem types.

In selecting independent variables, we noted that (i) diameter was most universally available, (ii) height was available in some cases, and (iii) specific gravity might be available when inventory data are reported by species and the species specific gravity is available. Although site index is another important and useful variable, we found that there was no consistency in the manner with which site index was reported. We developed three equations for each life zone grouping: a parabolic equation based on diameter alone and fit using weighted least squares (Cunia 1964), and logarithmic equations based on ($\text{diameter}^2*\text{height}$) and ($\text{diameter}^2*\text{height}*\text{specific gravity}$) using Baskerville's logarithmic coefficient adjustment (Baskerville 1972). These equations are discussed more fully in Brown et al. (1989).

The data used to develop these equations came from the many researchers in different parts of the tropics. The data were pooled within life zone grouping after testing for differences between samples. Although the differences were statistically significant in several cases, they were not large relative to the differences between life zone groupings and were ignored. There were important and statistically significant differences between different life zone groupings. Trees in the moist life zone tend to have more biomass than similar sized trees in wet or dry life zones. We developed a height-diameter relationship for the moist and wet life zone groupings using an extensive data set from Venezuela. This equation was combined with the biomass-diameter-height regression to provide double sampling regression equations with increased precision for inventories from similar forest types.

We use the estimated coefficient covariance matrix and conditional variance to estimate the error associated with the regression equation. Since most inventory stand tables are reported without the associated summary statistics necessary to estimate the error contribution of the stand table, the estimates of mean biomass per hectare only consider the error of the regression and are thus underestimates of the true error.

Ratio Estimation of Total Aboveground Biomass

Volume and biomass are functionally related: biomass is the product of the volume times the density or specific gravity, summed over all parts of the tree. The volume most commonly reported is some measure of commercial volume which ignores non-commercial portions such as branches, leaves, bark, and the tops of trees. If species is known, one can estimate commercial biomass from commercial volume by multiplying the volume by the specific gravity of the species. We use ratio estimation to estimate total aboveground biomass from commercial biomass.

We had several paired stand and stock tables by species, where the volumes had been calculated based on locally-derived volume equations generated as part of a formal forest inventory. We computed the commercial biomass per hectare, and estimated the total aboveground biomass by applying our regression equations to the reported stand table. The resulting estimates were tallied and the ratio of total to commercial biomass estimated by a mean-of-ratios estimator. Such an estimator would be useful in cases where stand tables are not available, but estimates of the total volume per hectare are available. These ratios can also be used to estimate biomass increments based on estimates of volume increment. When volume is not broken down by species, an average specific gravity for the forest type must be substituted (Brown et al. 1989).

There was great variation in the ratios among different forests. Brown and Lugo (1984) reported the same problem with their mean-of-ratios estimator. Inventory methods vary widely in the tropics, so the estimates of "commercial" volume are partially dependent upon the inventory methodology, e.g. minimum top diameters considered in estimating volume and different minimum diameter at breast height (dbh) sizes included in inventories. The sizes of the trees also make a difference: small trees have zero commercial volume up to some minimum merchantable limit, whereupon they suddenly have a positive commercial volume; then, as the tree increases in diameter, the proportion of the total volume counted as commercial also increases.

In trying to explain some of this variation, we plotted the ratio estimators against the quadratic stand diameter (Figure 1). Stands with a low QSD (15 cm) tend to have relatively high expansion factors, ranging in our study from 4.0 to 7.0. As the QSD increases from 15 cm to 40 cm, the expansion factor declines sharply in an exponential manner. Once QSD reaches 40 cm, the expansion factor appears to level off somewhere between 1.5 and 2.0. Once one accounts for differences in QSD, there is little variation in ratio estimator across life zones.

This relationship allows refined estimation of biomass in regions where little local information other than volume is available. Given basal area/ha and numbers of stems, one can compute QSD and then estimate an expansion factor to convert commercial biomass to total biomass. Primary undisturbed forests will tend to have higher QSD values, implying lower expansion factors. As forests become increasingly disturbed, the QSD is likely to decrease and the expansion factor to increase. Even when one has no information about QSD, one can at least narrow down the range of feasible expansion factors based on qualitative knowledge of the forest type (i.e. undisturbed vs. secondary forest vs. areas of shifting cultivation). We are currently conducting a sensitivity analysis of the ratio estimator to variables such as different minimum diameters of inclusion and different methods of volume estimation.

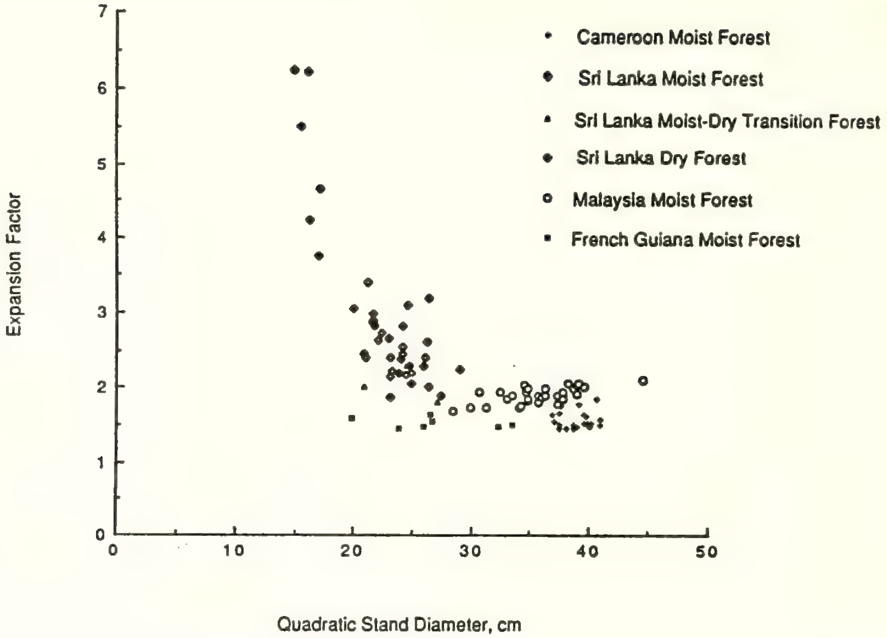


Figure 1. The relationship between the ratio of total to commercial biomass (Expansion Factor) and the Quadratic Stand Diameter based on inventories in six forest types from four tropical countries (from Brown et al., 1989).

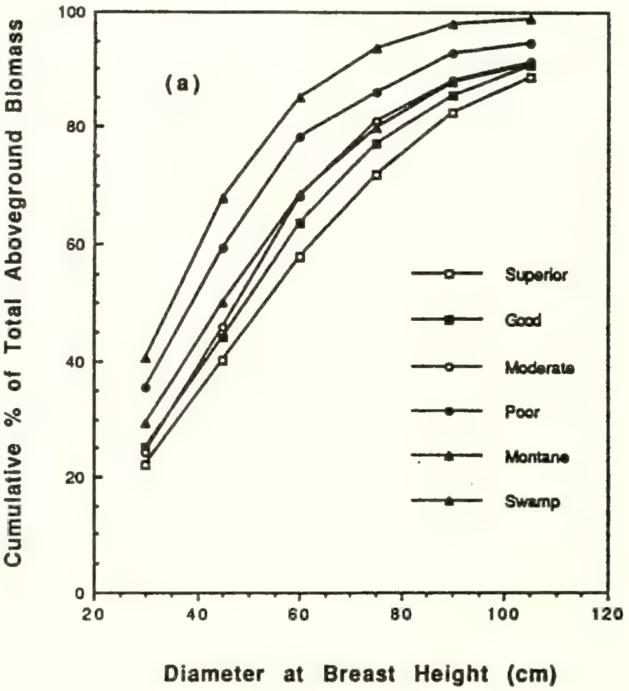


Figure 2. Cumulative distributions of total aboveground biomass by diameter class from six forest types in Peninsular Malaysia (from Gillespie et al., manuscript in preparation).

The Case of Incomplete Stand Tables

Part of our interest in carbon flows in the tropics requires that we deal not only with current data but with historical data as well, in order to attempt to estimate past fluxes of carbon due to land use change over the last century. Thus we work not only with current inventories but also with older inventories. Many of the older inventories tended to focus only on the largest, most commercially valuable trees, and to ignore trees under 30 or 40 cm dbh. We were concerned that our methods outlined above, based on reported information, would ignore a potentially important part of the total stand biomass when the stand tables omitted stems below a certain limit. In earlier work, we avoided this problem by considering only those inventories which included trees down to 10 or 15 cm dbh; stems below 10 cm are estimated to contain less than 3%-4% of the total stand biomass (Brown & Lugo 1984).

Since the current project required many more data points than those for which complete stand tables are available, we developed a method of estimating the biomass or volume in excluded diameter classes based on the volume or biomass in the present diameter classes. In order to accomplish this, we studied diameter distributions of tropical forests.

We began by computing the distributions of volume, basal area, and biomass by diameter class for a variety of forest types. The results indicated that trees below 35 cm dbh can contain from 25% to 45% of total aboveground biomass (Figure 2). This seemed sufficiently large to justify efforts to estimate the frequencies of the missing stems.

Most tropical forests tend to be uneven-aged and uneven-sized. Dawkins (1958) gave what he reported to be a "generic tropical diameter distribution" consisting of an inverse-J shaped curve, with many small stems and relatively few larger stems. When one plots the logarithm of the number of stems vs. the diameter class midpoint, the points often fall in approximately a straight line, implying that the ratio of numbers of stems in adjacent diameter classes is approximately constant over much of the diameter range. This ratio is often referred to as de Liocourt's q (Meyer 1952).

We used this approach to fit straight lines to a variety of stand tables from three tropical countries. The lines fit surprisingly well for all but the largest diameter classes in the undisturbed forests (R^2 ranging from 0.96 to 1.00 over 6 to 9 diameter classes). The slope of the fit line tended to increase as overall stand biomass or volume increased, which is expected since less disturbed forests have on average more of the larger trees. The q value appeared to remain constant over the smaller diameter classes, but generally decreased for the largest diameter classes.

Since we were interested in estimating stem frequencies of smaller diameter classes, we tested the hypothesis that a single q value might be used to estimate the frequencies of missing diameter classes. Using tables which were complete down to 10 or 15 cm, we dropped all diameter classes below 35 cm and tried several methods of estimating the missing diameter class frequencies. Our three approaches consisted of (i) determining q based on all remaining diameter classes, (ii) determining q based on all diameter classes with at least an average of 1 tree per hectare, and (iii) determining q by taking the ratio of the numbers of stems in the two smallest present diameter classes. We found that (iii) was the most reliable approach, with the estimates of total stand biomass within 10% of the estimates based on the complete known stand table (Gillespie et al., 1989). Since this error is less than that of simply ignoring missing stems, we feel comfortable in estimating the missing stems when no other information is available.

SUMMARY

Biomass estimates of tropical forests are important for a variety of management and research needs, yet there are relatively few in-place inventories which measure biomass directly. Some information can be obtained via remote sensing (e.g. Nelson and Krabill (1988)) but much of this methodology is still in the experimental stage and is of no use for evaluating historical changes in biomass. Since there is much information available dealing with forest volume estimates, and since volume and biomass are strongly correlated, it seems appropriate to estimate biomass based on volume. Given the wide disparity in the forest types and in the methods of collecting and reporting inventory information, no one approach will serve to yield estimates in all cases. A toolbox approach is called for, with a variety of estimation procedures documented and available for use as needed. We feel that the three methods outlined above are a good start in dealing with the problem. More tools are needed, especially in the areas of error estimation which can include an error component from the stand tables.

REFERENCES

Baskerville, G., 1972: Use of logarithmic regression in the estimation of plant biomass. Canadian Journal of Forest Research 2:49-53.

Brown, S., and Lugo, A. E., 1982: The storage and production of organic matter in tropical forests and their role in the global carbon cycle. Biotropica 14:161-187.

Brown, S., and Lugo, A. E., 1984: Biomass of tropical forests: a new estimate based on forest volumes. Science 223:1290-1293.

Brown, S., Gillespie, A. J. R., and Lugo, A. E., 1989. Biomass estimation methods for tropical forests with applications to forest inventory data. Forest Science, in press.

Cunia, T., 1964: Weighted least squares method and construction of volume tables. Forest Science 10:180-191.

Dawkins, H. C., 1958. The management of natural tropical high forest with special reference to Uganda. Institute Paper No. 34, Imperial Forest Institute, University of Oxford, Oxford, England. 155 pp.

Gillespie, A. J. R., Brown, S., and Lugo, A. E., 1989. Estimation of stem frequencies from truncated stand tables in three tropical moist forests. In preparation.

Holdridge, L., 1967: Life Zone Ecology. Tropical Science Center, San Jose, Costa Rica. 206 pp.

Houghton, R. A., Boone, R. D., Melillo, J. M., Palm, C. A., Woodwell, G. M., Myers, N., Moore, B., and Skole, D. M., 1983: Net flux of carbon dioxide from tropical forests in 1980. Nature 316:617-620.

Hozumi, K., Yoda, K., Kokawa, S., and Kira, T., 1969: Production ecology of rainforests in Cambodia, I. Plant biomass. Nature and Life in Southeast Asia 6:1-51.

Johnson, W. C., and Sharpe, D. M., 1983: The ratio of total to merchantable forest biomass and its application to the global carbon budget. Canadian Journal of Forest Research 13:372-383.

Meyer, H. A., 1952:Structure, growth and drain in balanced unevenaged forests. Journal of Forestry 50:85-92.

Nelson, R., Swift, R., and Krabill, W., 1988: Using airborne lasers to estimate forest canopy and stand characteristics. Journal of Forestry 86(10):31-38.

Ovington, J. D., and Olson, J. S., 1970: Biomass and chemical content of El Verde Lower Montane rainforest plants. In Odum, H. T. and Pigeon, R. F., eds., A Tropical Rainforest, Ch. H-2, US Atomic Energy Commission, Division of Technical Information, Springfield, VA.

Trabalka, J. R., (ed.), 1985: Atmospheric carbon dioxide and the global carbon cycle. United States Department of Energy, Washington, D.C. DOE/ER-0239. 315 pp.

Construction and Application of Biomass Regressions to Size-Classes of Trees[1]

Xiaoming Yang and Tiberius Cunia
Graduate Student and Professor of Statistics and Operations Research Respectively
State University of New York College of Environmental Science and Forestry
Syracuse NY 13210

ABSTRACT

It is common in forest inventory for regional biomass (volume) regressions (assumed unbiased for an entire population of trees from a given region) to be applied to inventories of specific forest areas (representing parts only of that region) or a subset (by size or utilization) of the forest trees. Using simulation techniques, it is shown that, proceeding this way, the bias may be negligible but the error due to the regression function may be unexpectedly high. It seems that a better procedure may be to construct new biomass regressions using those sample trees (used in the construction of the regional regressions) that are part of the population of trees of the forest area being currently inventoried.

INTRODUCTION

The biomass (volume) of a given forest area is generally estimated by a double (two-phase) sampling design. In the first phase, the trees of a relatively large sample are measured for (i) tree characteristics such as diameter, species, height, etc. and (ii) stand and ground attributes, such as forest type, age class, soil, site quality, etc. These trees are not measured for biomass, the main variable of interest. In the second phase, the trees of a relatively small sample are measured for biomass, in addition to some or all of the variables measured in the first phase trees. The second sample may be a subsample of the first or completely independent of it. The data of the second phase sample are used to estimate the relationship (usually a regression function) between biomass and the tree and stand attributes measured in the first phase. Applied to the data of the large, first phase sample, the relationship yields estimates of average biomass per acre by classes of stands or trees, or the corresponding totals for the entire forest area.

It is common to have a standard "phase 2" type of a sample of trees selected from a "large" forest area, usually a large geographical region. The data of this sample are used to construct regional biomass regressions, which applied to a variety of "phase 1" types of samples, selected from a variety of forest areas of interest, yield estimates of average or total biomass. More specifically the regional biomass regressions are applied, time and time again, to forest inventories of (i) the same or different forest areas (entire region or parts thereof), (ii) the same or at different points in time, or (iii) the same or differently defined classes of merchantable trees.

For example, a random sample of trees selected from forest areas of a given state, has been used to construct "regional" biomass regressions applicable to the entire state. These regressions are repeatedly applied to plots, periodically remeasured, of a statewide CFI (Continuous Forest Inventory) System. Or, they are applied to the biomass inventory of small forest areas. Or, they are applied to inventory of the biomass of the trees suitable for veneer production in a specific, small or large forest area, within or outside the state.

The biomass estimates have an error. Because of the structure of the sampling design, the error of the estimates has two main components; one due to the first phase, the others to the second phase sample. Each component in turn has also several subcomponents; one associated with the variation of the sample units, the second with the measurement of the units selected in a specific sample, the third with the statistical model used to derive the estimates and possibly others. Each subcomponent itself may have, in addition, a random error, which over the long run, tends to cancel itself out and a systematic error, a bias, which is largely unaffected by the sample size, or the number of times we sample. We shall be concerned in this paper with the error due to the second phase sample, (error of the biomass regressions), and we shall implicitly include in it, both random and systematic components.

---

[1] Paper based on a thesis submitted by the senior author in partial fulfillment of the requirements for a MSc degree at SUNY College of Environmental Science and Forestry, Syracuse NY.

Let us now consider, in more detail the four components of the error of biomass regressions. There is first the sampling error associated with the differences between the biomass of individual sample plots or trees. The same sampling procedure, applied on different occasions, results in different sets of sample trees and, thus, in different biomass regression functions.

The second component is the measurement error, defined here as the difference between the true and the recorded value of a sample unit measurement. The same or different people, using the same or different instruments and rules of measurement, end up ordinarily with different sample values. In the specific case of the measurement of the oven dry biomass weight of a sample tree, the error has an additional source, that of subsampling; the oven dry weight is determined, for practical purposes, by selecting a small sample from the tree, determining its oven dry weight and then extending that weight to the overall tree.

The third component is due to the error of statistical model used in the estimation of the regression function. Different statisticians working with the same sample data, may use different statistical models and, thus, arrive at different biomass regressions. For example, one may use linear or non-linear regression functions of different forms, may use the same or different estimation procedures (like least or weighted least squares), with the same or different independent variables, in their original form or some mathematical transformation, etc.

The last component is that of application error, defined here as the difference between the "true" regression functions of the population for which the biomass regressions was estimated and the one to which they are currently applied. Because we always deal with dynamic populations, there is always an application error, even when the biomass regression is applied to the "same" population, containing the same trees. This application error may be small, but sometimes it may be quite large. Although one may estimate this error, this is seldom done in forest inventory.

The objective of the paper is to have a look at the size of the error of application as it occurs in a particular case. More specifically, our objective is to (i) calculate a biomass regression from a sample selected from an entire forest area containing merchantable trees of all sizes, (ii) apply the biomass regression so calculated to subpopulations of forest trees of specific (not all) size classes, (iii) analyze the bias and precision of the estimates derived and (iv) compare the efficiency of this procedure to that obtained by the alternative procedure when in the calculation of the biomass regressions we use only the sample trees belonging to that subpopulation.

THE METHOD OF INVESTIGATION

To achieve the objective of the study, we have used simulation techniques. This involved several steps, from construction of forest tree populations to simulate, to selection of sampling and estimation procedures, to writing and application of computer programs and, finally, to the statistical analysis that leads to the final conclusions of the study.

Study Populations

The basic study population was constructed prior to the present research project and for more details the reader is referred to Cunia (1987). It suffices to say here that (i) the basic data used in the construction of the population consisted of measurements peformed on 22,753 trees distributed in 925 one-fifth acre plots of a forest inventory survey made some 30 years ago in the State of New York, (ii) the trees were measured in the field, among other things for their diameter (at breast height) and (merchantable) height and (iii) the total height and total above ground biomass (green weight) of each tree was generated by a Monte Carlo process which insured that the main characteristics of a real life population of forest trees (variation of biomass about regression function, within and between plots) were all preserved.

The original population is defined in this study as Population 1. Using its trees, we have defined five additional populations that were subsequently needed in our study, namely, (i) Populations 2 and 3 containing the trees of diameter (a) less than or equal to 10 inches and (b) greater than 10 inches respectively and (ii) Populations 4, 5 and 6 containing the trees of diameter (a) less than or equal to 7 inches, (b) larger than 7 but less than or equal to 12 inches and (c) larger than 12 inches respectively.

For each of the six populations we have calculated a set of parameters, the means per tree, the means per acre and the totals of several tree variables, all needed in the subsequent research. Among these variables are y = tree biomass, $x_1 = 1$ (a dummy variable), x_2 = tree diameter, x_3 = (squared diameter) @ basal area and x_4 = (squared diameter) (total height) @ volume, where "@" stands for "approximately proportional to".

The objective is to estimate, by sampling the mean biomass per acre. We have considered only "second phase" samples and their regressions were applied to the entire population. Thus, one main component of the error, that of the "first phase" sample has been eliminated because the sample is actually the population. The "measurement error" component has also been eliminated because the tree variables we had in the computer were assumed measured "without error". The only remaining error components were those of (i) sampling (second phase), (ii) statistical model and (iii) application as related to statistical models used.

Sampling Methods

The (second phase) sampling method is defined by Cunia (1987) as two-stage cluster sampling. In the first stage m clusters (sample plots) are selected from the 667 non-empty clusters (plots with at least one tree) of Population 1 by simple random sampling without replacement. In the second stage, p percent of the trees from each of the m selected clusters of the first stage were selected also by simple random sampling without replacement, and their biomass (y), diameter (d) and total height (h) recorded ("measured"). Not desiring to have samples that are too small or too large, we have used only the following combinations of m and p, (i) m = 20 and p = 30, 40, 50, 60, 70 and 80%, (ii) m = 30 with p = 20, 30, 40, 50% and (iii) m = 50 with p = 10, 20, 30%.

Each of the above 13 sample sizes (called here, for convenience, sampling methods) was applied repeatedly to Population 1 during previous research projects. Not having to do again the actual sampling process, we have simply used 100 sets of simulated samples already available for each of the 13 sampling methods above.

Because each of the 1300 samples so generated was to be applied to each of the six populations, each sample was re-defined as consisting of six new samples, one for each of Populations 1, 2, ... , 6. Thus, each sample provided six sets of biomass regresson functions, depending on whether all trees, or only trees of a certain size were used for the calculation of the regressions.

Estimation of Biomass Regressions

For each sample and each of the six populations we have used eight regression models, combinations of two forms of regression function and four estimation approaches. The two forms are:

Form 1: $y = B_1 + B_2 d + B_3 d^2 = B_1 x_1 + B_2 x_2 + B_3 x_3 = [B]'[x]$, and

Form 2: $y = B_1 + B_4 d^2 h = B_1 x_1 + B_4 x_4 = [B]'[x]$

where y, d, h, x_1, x_2, x_3 and x_4 have been previously defined.

Following are the four estimation approaches: Approach 1 - Ordinary Least Squares (OLS); Approach 2 - Ordinary Weighted Least Squares (OWLS) with the conditional variance of y assumed proportional to d^4 for Form 1 and d^4h^2 for Form 2; Approach 3 - Modified Least Squares (MLS); a procedure that takes into account the cluster effect of the sample data by working with the sum (over the sample clusters) of the tree variables not with the values of the individual trees; Approach 4 - Modified Least Squares (MWLS); a procedure similar to MLS but with the added assumption that the conditional variance of y is proportional to d^4 for Form 1 and d^4h^2 for Form 2. More details on these estimation approaches, and the reason for selecting the regression function Forms 1 and 2 can be found in Cunia (1987).

Estimation of the Mean Biomass Per Acre

The summary statistics of the second phase sample are (i) the vector [b], estimator of [B] and (ii) the matrix $[S_{bb}]$, estimator of the covariance matrix of [b]. It can be shown that the mean biomass per acre, say m_a, is equal to

$$m_a = B_1m_1 + B_2 m_2 + B_3 m_3 = [B]' [m]$$

and

$$m_a = B_1m_1 + B_4 m_4 = [B]' [m]$$

for the regressions of Form 1 and Form 2 respectively. In these equations m_i represents the (known) mean on a per acre basis of the variable x_i, i = 1,2,3,4 previously defined, and [ ] is the matrix notation we use in this paper.

As [b] is an estimate of [B], the statistics $z = [b]'[m]$ and $V = [m]'[S_{bb}][m]$ are estimates of m_a and variance of z respectively. If both [b] and $[S_{bb}]$ are unbiased, so are z and V. The bias of z, if any, can be estimated by the difference $d = (z-m_a)$. Recall that, contrary to what happens in real life, we know m_a.

For any given sampling method, estimation procedure and population, the 100 simulated samples provide up to 100 estimates z, say K estimates. Note that for some populations, some of the samples of trees may be too small to allow the estimation of the biomass regressions; in this case this sample is ignored. Using the K values z we have estimated (i) m_a, by the mean value $\bar{z}$ = (sum of z)/K, (ii) the bias of z, by the average difference $\bar{d} = \bar{z}-m_a$ and (iii) the variance of z, by the value S_{zz} = (sum of $(z-\bar{z})^2$)/(K-1).

Note that both V and S_{zz} estimate the variance of z. However, V is a sample-based estimate, that is, it uses the information from one sample only (and, thus, is always available) while the calculation of S_{zz} requires several samples (something which rarely occurs in real life). The advantage of S_{zz} is that it is a pure estimate of the variance of z and it takes into account the effect of sampling method and estimation procedure, something which is not true for V.

The null hypothesis that the bias is zero can be tested by the statistic $t = (\bar{z}-m_a)/\sqrt{S_{zz}/K}$ which has the t-distribution with (K-1) degrees of freedom. We do not have a good statistical test to compare S_{zz} to $\bar{V}$ = (sum of V)/K. However, assuming statistical independence of S_{zz} and $\bar{V}$ we may use the approximate F test applied to the ratio $\bar{V}/S_{zz}$. Because $\bar{V}$ and S_{zz} are not independent, however, all sample values F must be interpreted with extreme care.

To estimate m_a of Population 1, one would normally use the biomass regressions calculated from the entire sample of trees of all sizes. To estimate m_a of the other five populations, however, one can use either (i) the regressions of Population 1, or (ii) the biomass regressions calculated from the part of the sample that includes only the trees of the size belonging to the given population. For example, the mean m_a of Population 3 can be estimated either by the biomass regressions of Population 1 or those of Population 3, that is, the regressions which have been calculated only from the sample trees that have a diameter greater than 10 inches. Recall that Population 3 is a part of Population 1.

For any given sample, and if i = 1,2, ... ,6 denotes the population number, we have, for each $i \neq 1$, (1) $[b^1]$ and $[b^i]$, two estimates of [B], (2) $[S^1_{bb}]$ and $[S^i_{bb}]$, estimates of the covariance matrices of

$[b^1]$ and $[b^i]$ respectively, (3) $z^1 = [b^1]'[m]$ and $z^i = [b^i]'[m]$, estimates of the sample mean m^i_a, (4) V^1 and V^i, sample-based estimates of the variances of z^1 and z^i respectively.

For any given set of 100 (or K) samples, we have, for each $i \neq 1$, (1) $\bar{z}^1$ and $\bar{z}^i$, two estimates of the same mean m^i_a, (2) S^1_{zz} and S^i_{zz}, estimates of the variances of z^1 and z^i respectively, (3) $\bar{d}^1$ and $\bar{d}^i$, estimates of the bias, if any of z^1 and z^i respectively, (4) $t^1 = \bar{d}^1/\sqrt{S^1_{zz}/K}$ and $t^i = \bar{d}^i/\sqrt{S^i_{zz}/K}$, the value of the test statistic for the null hypothesis that the bias of z^1 and z^i respectively is equal to zero and (5) S^1_{zz}/S^i_{zz}, a measure of the ratio of the variances of the two estimators z^1 and z^i of the m^i_a, calculated by the two sets of biomass regressions.

Numerous tables were constructed to list many statistics by (i) sampling method (thirteen combinations of m and p values), (ii) biomass regression model (eight combinations of two function forms and four estimation approaches), (iii) set of sample tree data used to estimate the regressions (either data of all tree sizes or data of trees belonging only to a given population) and (iv) population number. Some of the statistics listed in these tables are: (1) the bias (in absolute value or percent of the mean) and the value of the t-test statistic for the null hypothesis that the bias is equal to zero, (2) the standard errors of the estimates z of m together with (i) the ratio of the two variances S^1_{zz}/S^i_{zz} and (ii) the F-value of the approximate test of the null hypothesis that z^1 and z^i are equally precise, (3) standard error of the average $\bar{V}$ of the sample-based estimates of the variances of z, together with the ratios of the variances $\bar{V}^1/\bar{V}^i$ and $\bar{V}/S_{zz}$, and (4) the average differences $\bar{d}^i = \bar{z}^1 - \bar{z}^i$ and the values t of the test statistic for the null hypothesis that the expected value of d^i is equal to zero.

ANALYSIS OF RESULTS

An ocular analysis of some 100 tables leads to the following main conclusions:

(1) With very few exceptions, the bias of the estimates z^1 and z^i, for all eight regression models, all sampling methods and all populations, is not significantly different than zero. In relative terms, the sample values of the bias are ordinarily less than 1% and, in general, it is smaller for z^i than the corresponding z^1. The few exceptions are probably due to the fact that a very large number of test statistics have been considered in the study.

(2) For the ordinary (OLS and OWLS) regression models, the precision of z^1 is (i) slightly better than that of z^i (although not significantly better) in about 50% of the cases and (ii) not as good as the precision of z^i (sometimes significantly worse) for the other 50%. This seems to imply that if the ordinary least and weighted least squares method is being used, one should probably use z^i rather than z^1. For the modified (MLS and MWLS) regression models there seems to be no doubt; the variances of the estimators z^1 are ordinarily between 2 and 6 times higher than those of estimators z^i, sometimes as large as 13 times.

(3) The statistic V is the estimator of the variance of z under the basic assumptions of the regression model used. This is the only estimator of the variance of z available in real life and it is valid only when the basic assumptions are strictly satisfied. Cunia (1987) has shown, in the past that, for Population 1 (where no sample tree is from outside the population) the variance of z is (i) grossly underestimated by the OLS and OWLS regression models and (ii) slightly overestimated by the MLS and MWLS models. These conclusions are expected to hold true for the estimators z^i (where the sample trees used in the calculation of the biomass regressions are selected only from the trees of the corresponding population i). A look at the results of the present study shows that the variance of z^i is (i) as expected, grossly underestimated for the OLS and OWLS models and (ii) about the same or slightly overestimated about 50 percent of the time and underestimated (between 10 and 20 percent) the other 50% of the time, for the MLS and MWLS models. When estimators z^1 rather than z^i are used, the results fluctuate wildly with no discernible pattern. The precision may sometimes be adequately estimated, but most of the time it is

not. The sample based variance V of z^1 may be anywhere between less than .10 and 1.30 times the value as estimated by the corresponding S_{zz}.

FINAL RECOMMENDATIONS

In view of the above conclusions, what recommendations can we give to the forest manager? He is faced with the problem of taking the inventory of the forest trees of a given size class and the biomass regressions he has, have all been calculated from trees of all sizes, selected from the area being currently inventoried or a similar forest area. We may distinguish several possible cases.

Case 1 - The biomass regressions have no adequately expressed statement of error associated with them and the sample trees from which they were constructed are not available. Then, the manager may have to either (i) use the regressions as they are, but then he should be aware that the error may be much larger than expected or (ii) select a new set of sample trees and construct new biomass regressions.

Case 2 - The trees from which the biomass regressions have been calculated are available, but the subsample of trees that may be considered as part of the tree population of the area being currently inventoried is too small to be of any use in making new, local biomass tables. Then, the manager finds himself in the same situation as in Case 1 above.

Case 3 - Same as Case 2 but the subsample of trees is sufficiently large. Then, our advise is to calculate new biomass regression functions, to estimate their error and include this component in the final inferences about the error of the biomass estimates. It is our feeling that proceeding this way, one would reduce the chances of having a large error component associated with the biomass regressions.

In all three cases, the bias of the all-size biomass regressions would be small, if any, provided that these regressions are indeed applicable to the forest area of interest.

A general recommendation we feel may be worthwhile considering for the future is that of replacing the use of regional biomass regressions by the use of large data sets of sample trees selected at random from a relatively large region. Whenever a "phase one" type of sample of trees is selected from a given forest area of interest being currently inventoried, the computer program that calculates the biomass statistics should have a subroutine which (i) selects (from this sample tree data set) the sample trees that have sizes and characteristics similar to those of the forest area of interest, (ii) constructs new biomass regressions and calculates their error and (iii) adds this error to the error of the "first phase" sample, when estimates of the forest area biomass are being made.

REFERENCES

Cunia, T., 1987: Evaluating errors of tree biomass regressions by simulation. In: Wharton, H.E. and Cunia, T. Eds. - Estimating tree biomass regressions and their error, Publication Nr. NE-GTR-117, USDA Forest Service, Northeastern Forest Experiment Station, Broomall PA.

FRACTAL GEOMETRY AND FOREST MEASUREMENTS

Boris Zeide
Professor of Forestry
Department of Forest Resources, University of Arkansas at Monticello,
Agricultural Experiment Station, Monticello, AR 71655

ABSTRACT

Virtually all forest measurements are based on the integer dimensions of classical geometry. It is believed that a given tree can be described by a unique length or area independent from the unit of measurement. Fractal geometry calls for a major revision of this concept. Unlike the straight lines of classical geometry, natural lines do not have unique invariable lengths. They become longer when we use smaller yardsticks (units of measurement) to measure them. The length of natural lines is a process rather than an event. This process is controlled by a constant parameter which is unique for each line. This parameter can be identified as a spatial dimension. This dimension exceeds the corresponding integer dimension. It is fractional and is, therefore, called fractal dimension. Natural lines are hybrids between ideal lines and surfaces, just as classical surfaces and volumes serve as the ideal extremes of actual surfaces. Application of these novel concepts to forestry requires the development of methods suitable for trees. Two promising methods for determining fractal dimensions of crown surface are outlined. At present, the application and development of fractal geometry appears to be the most promising and profound direction of forestry research.

HOW TO MEASURE NATURAL OBJECTS

We have reached a high degree of sophistication in measuring natural objects such as trees and their parts. Numerous methods and techniques described in forest mensuration textbooks testify to the advanced level of our knowledge. Yet, some basic problems of our trade remain unsolved. How, for example, would we measure crown surface?

I talked with a researcher who was taking photographs of trees and separate branches in order to determine crown surface. He was going to give these pictures to specialists in image processing who would digitize the pictures on an electronic scanner and, he assured me, measure the surface with high precision on a computer. But what exactly are they going to measure precisely? Should we measure the surface of a convex hull that envelopes the crown or should we take into account cavities, cuts, and gaps of the crown? If so, what is the smallest cavity which should be considered? Should foliage area be used as crown surface? Should we pay attention to the ups and downs of foliage surface? Or should we, perhaps, go deeper and use the surface of chlorophyll cells or their photosynthetic components? Even if we disregard the intricacy of crown structure, it will not be easy to characterize the surface of these components. After many frustrating attempts to grasp the complexity of molecular surfaces, the physicist Wolfgang Pauli said that "the surface was invented by the devil" (Peterson 1989, p. 200). Without specifying the measurement unit, it is pointless to talk about crown surface.

Computer specialists are not in a position to answer these questions. It is our responsibility to tell them what to measure. Yet, although tree crowns have been the subject of much intensive research, we do not know how to measure the surface which is vital to tree existence. We are left to wonder how, without knowing the surface area, we can measure tree respiration per unit area, productivity, and dozens of other variables crucial to tree functioning? How can we study trees at all?

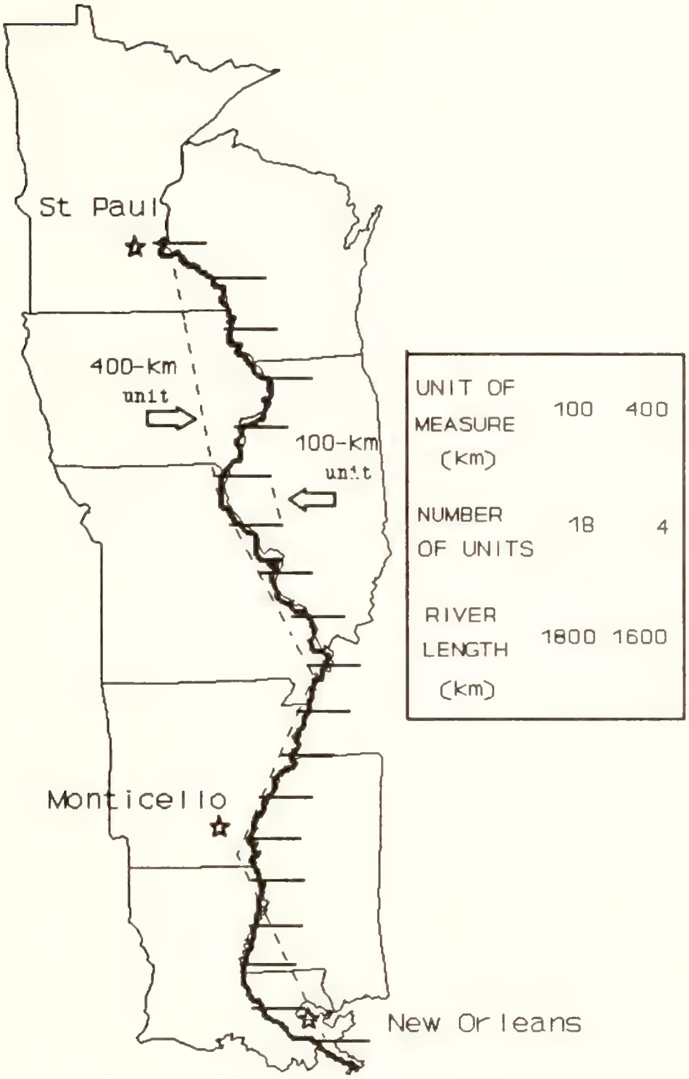


Figure 1. Unit size and corresponding length of the Mississippi river.

Similar problems appear when we measure the length of natural objects such as a boundary between forest types, a mountain outline, a river bank, or a coastline. Unlike the straight lines of classical geometry, natural lines do not have unique invariable lengths. Their lengths depend on the size of the unit of measurement, as is the case with the length of the Mississippi river (Figure 1). Without the unit size, the length alone makes little sense. Natural lines become longer when we use smaller yardsticks (units of measurement) to measure them.

Because the choice of the unit is arbitrary, the length might appear arbitrary too. Such uncertainty would make measuring useless. Fortunately, there is strict correspondence between unit and length. The existence of this relationship restores our trust in measurement. We trust, however, not length alone but the relationship of length to unit. The length of natural lines becomes a process rather than an event. This process is controlled by a constant parameter which is unique for each line. This parameter is called fractal dimension.

FRACTAL DIMENSIONS

Spatial dimensions, we were taught, are whole numbers (integers). Lines are measured by their length and have one dimension (10 linear feet, for example). Areas have two dimensions because they are measured by length and width. For instance, an area can be 10 x 12 = 120 square feet. And volumes have all three dimensions, length, width, and height. That is why we measure volume in cubic units. Length, areas, and volumes of ideal objects (such as straight lines, rectangles, and cylinders, or their approximations which we see in our buildings, sawn timber, and most man-made things) are independent from the unit of measurement. For a given object, the product of the unit size (u) and the number of the units (n), that is, length (c) is constant:

$$nu = c; \quad n = cu^{-1} \tag{1}$$

The power of u in the last equation (taken with the opposite sign) can be identified as the dimension of a line. On a plane where the unit size is plotted on the x-axis and the number on the y-axis, all geometrical objects are presented by three lines (Figure 2).

The rest of this plane is not as empty as implied by classical geometry. When the length is a function of the unit, as we saw with natural objects (Figure 1), its dimension becomes a fraction, for example, 1.08, which is specific for each line. The excess over one (1.08 - 1 = 0.08) shows how twisted or curvy a line is. Similarly, the dimensions of natural surfaces (such as the surface of a tree crown) exceed two. Natural lines are hybrids between ideal lines and surfaces, just as classical surfaces and volumes serve as the ideal extremes of actual surfaces. These natural objects fill the rest of the plane.

The relationship between the number of units and their size can be derived from the observation that natural objects with their seemingly irregular forms often have the simplest regularity: the pattern of curvature remains invariant and is consistently repeated on smaller and smaller scales. In creating diversity nature uses the same trick over and over again. Under magnification each segment of a line appears roughly the same as the entire line. Such lines are known as fractal lines. Though simple, this self-similarity, or scale invariance, can produce an astonishing variety of forms. The problem is finding a concise way to describe this pattern and, thus, to describe length. As we will see, the key to the solution is to revise the Euclidean idea of integer dimensions.

Let's construct an example of a fractal line and calculate its dimension. Take a straight line of unit length (c = 1), divide it into three equal sections, draw an equilateral triangle on the middle section, and remove the base of the triangle. The transformed line is shown in Step 1 (Table 1). Repeat this step for each of four straight segments (step 2) and continue the same procedure for the resulting segments. In Table 1 the number of units is presented as a function of the number of steps

$$n(k) = 4^{k} \tag{2}$$

To compare the number of units in the thus constructed curve (called the Koch curve) with the number of units in the Euclidean line (equation 1), we have to present the number as a function of the unit size, u. To do this let's express k of equation (2) as a function of u. Because $u = (1/3)^{k}$

$$k = -\ln u/\ln 3 \tag{3}$$

where ln is a natural logarithmic function. Therefore,

$$n(u) = (4^{\ln u})^{-1/\ln 3} \tag{4}$$

Because $4^{\ln u} = u^{\ln 4}$, the number is equal

$$n(u) = u^{-\ln 4/\ln 3} = u^{-1.26} \tag{5}$$

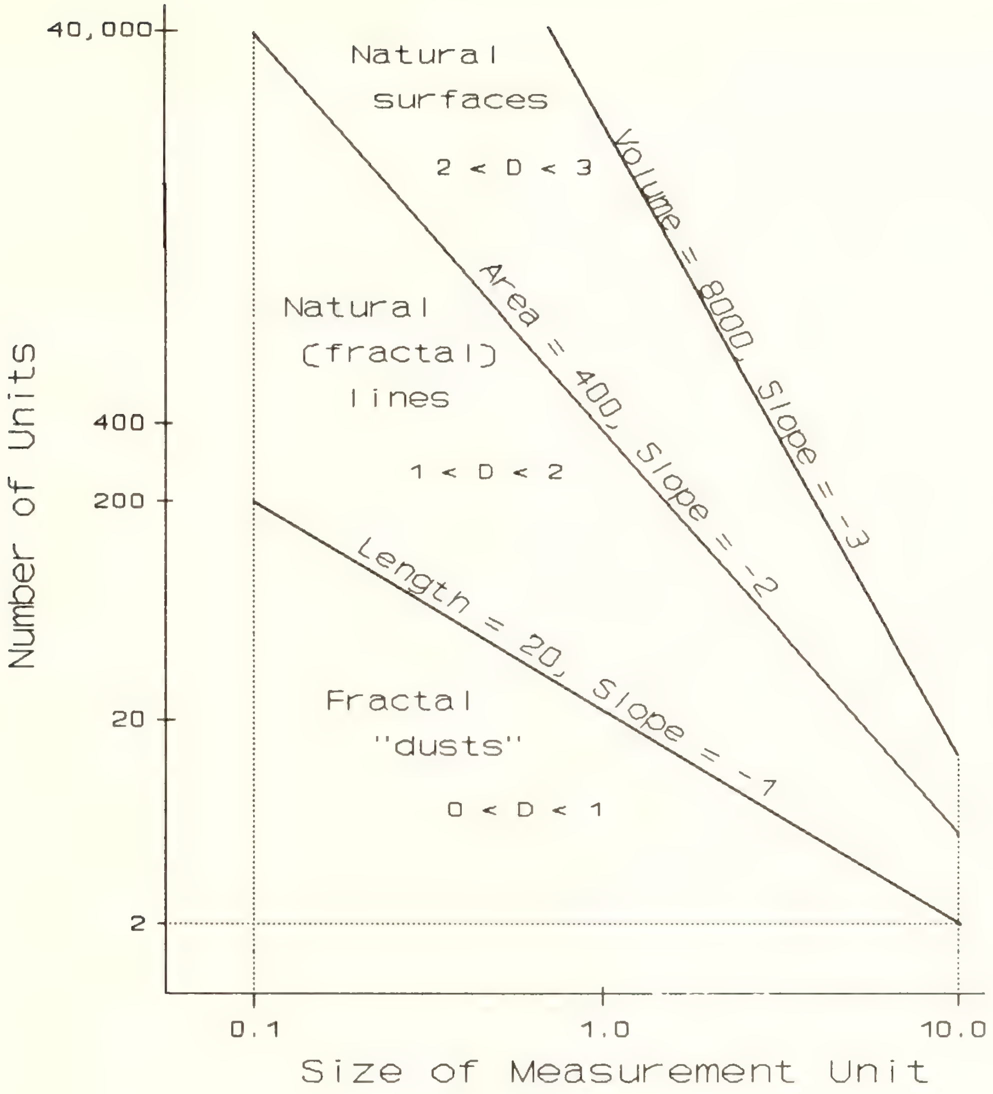


Figure 2. Lines and surfaces in classical geometry and in nature.

and length

$$L = un = uu^{-1.26} = u^{-0.26} \tag{6}$$

or, in general,

$$n = cu^{-D} \tag{7}$$

and

$$L = un = cu^{1-D} \tag{8}$$

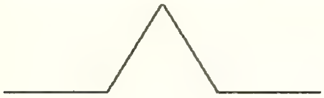
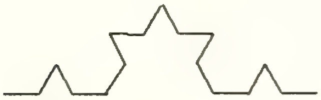
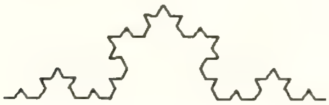
where D is a fractal dimension. For example, when the unit size is 1/3 (first step),

$$n(u) = 1/3^{-1.26} = 3^{1.26} = 4 \quad \text{and} \quad L = (1/3)4 \tag{9}$$

Now, we are able to calculate the fractal dimension of any line. From equation (7) and the measured lengths of Mississippi river (Figure 1) it follows that its fractal dimension is

$$D = \ln(18/4)/\ln(400/100) = 1.085 \tag{10}$$

Table 1. Dimensions of natural objects.

Line	Step	Unit size	Number of units	Length
	1	1/3	4	(1/3)4
	2	$(1/3)^2$	4^2	$(1/3)^2 4^2$
	3	$(1/3)^3$	4^3	$(1/3)^3 4^3$
	k	$(1/3)^k$	4^k	$(1/3)^k 4^k$

Because the power of equations (1), (5), and (7) was identified as a spatial dimension, we have to recognize the existence of fractional, or fractal, dimensions. These calculations are not an algebraic trick. They reveal that the description of nature requires a new geometry called fractal geometry (Mandelbrot 1983).

ESTIMATION OF FRACTAL DIMENSIONS OF CROWN SURFACE

The warped and discontinuous shapes of tree crowns suggest that fractals might provide a more meaningful and useful picture than does Euclidean geometry. Conceptually, the determination of fractal dimension of a surface is similar to that of a line. When natural surfaces are "paved" with successively smaller units of area, u^2, their areas, A, predictably increase:

$$n = cu^{-D} \quad \text{and} \quad A = u^2 n = cu^{2-D} \quad (2<D<3) \tag{11}$$

In so-called mass fractals, volume becomes indistinguishable from surface.

Practically, however, measuring surface dimensions is more complicated. The standard method for determining fractal dimensions, the box-counting method, is hardly suitable for trees. This method requires slicing a given object into many layers without any distortion of its structure, subdividing them into cubic boxes, and counting the number of nonempty boxes. This procedure is repeated many times using various sizes of the boxes. Fractal dimension of the studied object is one of the parameters of the relationship between the number of boxes and their size. Clearly, this method could be applied only on a limited scale in studies of tree crowns. To launch the study of fractal geometry of trees, we have to find another way of determining their dimensions.

One of the possible methods for determining fractal dimensions of tree crowns was developed in 1988 during my sabbatical in Finland. This method is called "the two-surface method" because it is based on the assumption that the relationship between leaf area of a tree, A, and surface area of a convex hull that envelopes its crown, E, contains a parameter which can be identified as a fractal dimension of crown surface, D. This relationship is

$$A = aE^{D/2} \tag{12}$$

where a is a constant. Because procedures for measuring the two required surfaces are well-established in biomass studies, the proposed method presents no technical problems.

The results produced by the two-surface method seem to be sensible. The slope of the regression of the logarithm of leaf mass (which was assumed to be proportional to leaf area) on the logarithm of the cubic root of crown volume (which represents the linear size of the crown) was calculated using published data for dominant trees of 10 conifer species measured in the Rocky Mountains (Brown 1978). Intermediate trees were measured for four species. The regression (on the log-log scale) was linear. This shows that crown surface is self-similar. The slope of the regression represents a fractal dimension. It varied between two and three (more precisely, from 2.27 for intolerant ponderosa pines to 2.81 for very tolerant hemlocks), that is, within theoretical limits of two and three. A value of three indicates that needles were uniformly distributed throughout the crown volume and their mass is proportional to the volume. This would be possible only if needles were extremely tolerant to low light. A dimension of two could be realized in the opposite situation when needles would not tolerate any shading and would be located exclusively on the crown periphery (so that their mass and surface are proportional to the surface of the hull). For these reasons it is not surprising that correlation between fractal dimensions and species tolerance was detected.

Usually, two-dimensional projections of a crown (or any mass fractal), such as a photograph, can be used to calculate the dimension of crown face ("the sieve dimension") rather than the dimension of the crown itself ("the sponge dimension"). The sieve dimension, D', provides little information about the sponge dimension, D. These dimensions are related only by an inequality (Pfeifer and Avnir 1983):

$$D' + 1 > D \tag{13}$$

Recently, Pfeifer (1988) proved a theorem that makes it possible to determine the fractal dimension of a spatial object, such as a tree crown, from two-dimensional projections. He showed that the distribution of holes, as they appear on the projection, still reflects the dimension of the original (unprojected) object. Analysis of this distribution allows one to calculate fractal dimensions. Although this method has not yet been tested on trees, it appears to be very attractive to forestry research because it involves fewer assumptions than the two-surface method, would require little field work, and potentially (using a combination of a video camera and a computer) could produce an instant answer in the field for each selected tree.

DISCUSSION

It is likely that fractal geometry will become an integral part of forestry methods because this new science provides concepts and tools to describe natural objects like trees. Besides mere description, fractals might facilitate our understanding of tree functioning. Because crown structure is related to many tree characteristics, both internal (such as tolerance or longevity) and external (such as site quality and stand density), fractal dimensions of the crown might find an explanation in terms of physiological and environmental factors.

In addition to its application to tree morphology, the fractal approach can contribute to the understanding of growth and development of forest stands because this approach plays the central role in studies of the dynamics of any complex system. The solution to the problem of succession might be a certain iterative function with parameters which can be interpreted in terms of environmental factors and properties of tree species. What makes the fractal approach particularly promising is that it brings together seemingly incompatible things such as chaos and order, ideal and real, reductionism and holism, finite and infinite, one and many, repetition and novelty.

Fractal geometry is still a young science and we should not expect that it would solve all the problems of forest measurement. Real trees are much more irregular and complex than the computer pictures that adorn publications on fractal geometry. Not all forestry patterns can be described by fractals. For example, according to Mandelbrot (1983, p. 161b), the ratio (tree height)^3/(stem diameter)^2 is constant. This means that stem profile is concave. Actually, it is convex. The 3/2 power law of self-thinning fits neatly into fractal geometry; yet, it was shown that the power function distorts the relationship between the number of plants and their average mass (Zeide 1987). Also, it is necessary to rememeber that not all tree variables are fractals. Stem surface is fractal but volume is not. Therefore, volume should be calculated using clsssical geometry. The same is true for tree height since straight stems provide a selective advantage in competition for light.

These examples are cited here not to discredit fractal geometry but to indicate that this science does not solve all forestry problems without our involvement. So far we have been mostly passive observers of genuinely significant scientific achievements, as with physics and cosmos research, or at best consumers, as with the computer revolution and molecular biology. Fractal geometry is the first area of fundamental research where we, as foresters, can actively participate rather than watch and wait for results obtained by others. Unlike the geometry of ideal objects which is the realm of pure mathematics, shapes of natural objects like tree crown should be studied jointly by mathematicians and ecologists to the mutual benefit of both disciplines. At present, the application and development of fractal geometry appears to be the most promising and profound direction of forestry research.

ACKNOWLEDGMENTS

I am grateful to Dr. Mandelbrot, the originator of fractal geometry, for his valuable advice on measuring fractal dimensions of trees. He also put me in touch with Dr. Peter Pfeifer who made available his works in advance of publication and shared a computer program that implements his promising method. My work was greatly encouraged by the enthusiastic response to fractal geometry from many foresters. In particular, I would like thank Dr. Seppo Kellomaki and Dr. Risto Sievanen who arranged my sabbatical work in Finland and my colleagues at the University of Arkansas. Valuable comments on this paper by associates Daniel Leduc and Suzanne Wiley are highly appreciated.

REFERENCES

Brown, J. K. 1978. Weight and density of crowns of Rocky Mountain conifers. USDA For. Serv. Res. Pap. INT-197. 56 p.

Mandelbrot, B. B. 1983. The fractal geometry of nature. W. N. Freeman, New York. 468 p.

Peterson, I. 1989. Scanning the surface. Science News 135:200-203.

Pfeifer, P. 1988. Fractals in surface science: scattering and thermodynamics of adsorbed films. Pages 283-305 in Chemistry and physics of solid surfaces, volume VII, R. Vanselow and R. F. Howe (editors). Springer, Berlin.

Pfeifer, P., and D. Avnir. 1983. Chemistry in noninteger dimensions between two and three. I. Fractal theory of heterogeneous surfaces. J. Chem. Phys. 79:3558-3564.

Zeide, B. 1987. Analysis of the 3/2 power law of self-thinning. Forest Science 33:517-537.

Remote Sensing and Forest Inventories

Moderator: Alan R. Ek

M.J. Duggin, P.F. Hopkins, and R.H. Brock, A Survey of Remote Sensing Methodology for Forest Inventory

Lawrence R. Gering, Dennis M. May, and Kurt B. Teuber, The Use of Aerial Photographs and Angle-Gauge Sampling of Tree Crown Diameters for Forest Inventory

M.E. Bauer, T.E. Burk, A.R. Ek, S.C. Ahearn, and L.P. Queen, Satellite Remote Sensing and Inventory of Minnesota Forest Resources

Paul Biggs, Inventory of the Jarrah Forest of Western Australia Using Large-Scale Photography and Ground Sampling

David Chung Liang Lee, René Antonio Novaes, Pedro Hernandez Filho, and Flávio Jorge Ponzoni, A Practical Example of Forest Mapping and Timber Volume Estimation Using Landsat-5 TM Imagery

Rachel Riemann Hershey, An Aerial Photographic Key as an Aid to Improving Photointerpretation Accuracy

G. Deegan and W.A. Befort, Photointerpretation Accuracy Across Two Decades of Forest Inventory

C.J. Liu, On Using Non-Imageable Point for Measuring Object Height on Vertical Aerial Photograph

W.A. Befort and G. Deegan, Organizing Statewide 35mm Photography in Support of Forest Inventory

Josef Hackl and Erwin Knappitsch, A Study of Pinus Mugo-Stands in the Karwendel Mountain Region (Tyrol, Austria)

A Survey of Remote Sensing Methodology for Forest Inventory

M.J. Duggin, Professor
P.F. Hopkins, Assistant Professor
R.H. Brock, Chairman and Professor
State University of New York
College of Environmental Science and Forestry
Syracuse, New York 13210

ABSTRACT

It has been 50-60 years since remote sensing (photographic) techniques were first used in forest inventory. A brief history of this activity is given leading into a discussion and review of established remote sensing activities in forest inventory. These include forest classification, forest mensuration procedures and sample designs for forest inventory. Current developments in forest inventory are discussed which cover digital image acquisition and analysis, video imagery, radar imagery, analytical photogrammetry, laser altimetry, geographic information systems and spatial positioning. The authors' thoughts are expressed concerning analog and digital image acquisition and analysis. A look to the future in imaging spectrometry is presented which includes a table of spectral regions for the remote sensing of vegetation. It is concluded that, although some perceive modern remote sensing as a novel, complicated and unproven tool for forestry applications, a cautious acceptance of emerging techniques can improve the amount of reliable information available to present day resource managers and inventory specialists. The extensive reference list includes publications dealing with remote sensing for forest inventory and related topics.

INTRODUCTION

Overview of Remote Sensing in Forestry

Remote sensing is used in many aspects of modern forestry, including classification and inventory, damage assessment, management (harvest planning and monitoring, reforestation), and recreation (Heller and Ulliman, 1983). In fact, forest type mapping and damage assessment were among the first remote sensing applications, when aerial photographs became widely used in the 1930's and 1940's. This paper reviews the use of remote sensing in forest inventory, emphasizing estimation of amounts of woody fiber.

Importance of Remote Sensing in Forest Inventory

An indication of the importance of forest inventory as a potential application of remote sensing technology may be obtained from the statistics of the Canadian Remote Sensing On-line Retrieval System (RESORS), which is probably the most comprehensive remote sensing data base in existence. Out of the 67,477 documents listed, 1,639 deal with forest inventory, 1,219 occurrences relate to forestry, and 655 relate to forest damage. While there will be considerable overlap between these categories, we obtain some idea of the considerable interest in applying remote sensing to forestry studies.

Forest inventory includes forest classification and forest volume estimation. Remote sensing techniques, especially aerial photographs, are used in both of these activities of forest inventory. In fact, remote sensing is used operationally in practically every unit which deals with forest inventory. Aerial photographs or other images are very commonly applied during forest classification (Heller and Ulliman, 1983). In addition, a variety of established techniques are available to precisely estimate the volume of wood resources on practically any sized area using aerial photographs, statistics, and appropriate ground reference information (Aldrich et al., 1984; Husch et al., 1972; Paine, 1981; Spurr, 1948). These forest mensuration activities using imagery are probably best described as applied photogrammetry. Inventories can be accomplished without aerial photographs or other remote sensing procedures, but usually with lower efficiency and greater restrictions for the same precision.

Brief History

In the U.S., aerial photographs were first used for forest inventory in the 1930s. Aerial photographs in forestry became more common in the late 1940s, especially after the publications by Spurr (e.g. Spurr, 1948). Many refinements and evaluations of these remote sensing techniques occurred in the 1950s and 1960s. Starting in the 1960s, new image types and advanced analytical procedures appeared. Since then, many experiments and improvements have been reported and some have been validated operationally. However, the complexity, expense, and novelty of these procedures has slowed their acceptance in forest inventory.

ESTABLISHED REMOTE SENSING ACTIVITIES IN FOREST INVENTORY

Forest inventory is accomplished with a number of activities which range from preliminary planning to information management to final communication about the amount of the resource available. Remote sensing may play a role in several of these activities. However, remote sensing has greatest potential in two phases of the process: forest classification and forest mensuration. The way in which remote sensing is used in these two phases will help determine and implement the inventory sampling design.

Forest Classification

Forest classification deals with identifying and usually mapping homogeneous areas (stands) of forest land. Often this activity is also termed forest type mapping, although this latter term usually infers a high degree of specificity in forest stand description. Remote sensing is the normal method by which forest classification is accomplished.

Classic Procedures. Classic procedures in forest classification use visual interpretation of aerial photographs to segregate forest regions into uniform areas. These uniform areas may be simple stratifications of forest or detailed descriptions of forest stands. Many references from the 1940s, 1950s and 1960s detail specific applications. The techniques are described generally in publications beginning with Spurr (1948) and Moessner (1953) and continuing to current times (Aldrich, et al., 1984; Avery, 1978; Burkhart, et al., 1984; Heller and Ulliman, 1983; Paine, 1981). Such methods are well understood and established. Typically, medium scale (1:10,000 to about 1:24,000), large format (9" by 9"), black and white photography is used. Information about the type of forest is usually supplemented by information about tree size (usually height), and stand density. These classes are mapped and used for many subsequent management activities. In forest inventory, the classes often provide strata useful for further sampling of the forest resource with larger scale aerial photographs and ground measurements. Greater detail in the forest classes requires larger scale photography. Species determination of individual trees is possible with very large scale imagery (Sayn-Wittgenstein, 1960). Photointerpretation keys have been developed to assist in classifying forest species or types (e.g. Befort, 1986; Tueller, 1979;

Sayn-Wittgenstein, 1978). Very large scale photographs were difficult to obtain with high quality with the early cameras and techniques, unless using helicopters.

Recent Developments. The first significant enhancement to these established techniques was provided by newer film types (first black and white infrared and then normal color and color infrared). The classification process became more effective because differences between coniferous and broad-leaved species were emphasized and more information was provided about subtle type differences. Use of the color film types is merely an extension of the established techniques with the added (and important) dimension of color. As presented by Aldrich (1979), color and especially color infrared films are more efficient and accurate for forest classification. As a result, smaller scale color photography usually provides classification information which is equivalent to interpretations from larger scale black and white film (Heller and Ulliman, 1983).

Other advances have also now improved the applicability of aerial photography to forest classification. Some of the significant advances relate to acquisition of very large scale photographs, use of inexpensive small format cameras, and the introduction of small scale aerial photography. Another significant development in forest classification is the general availability of digital satellite imagery and image processing.

Large scale aerial photographs are now easily acquired using small format (usually 70 mm.) cameras which are readily available with excellent capability and quality. Two general directions may be taken when using large scale photography. In the first, photographs are taken with either 35 or 70 mm. cameras and interpreted for species, density, or condition information. Scales are usually in the range of 1:3000 to 1:10,000, although even larger scales are possible (Needham and Smith, 1987). Determination of the exact scale of the photographs is unimportant since precise measurements are not taken and mapping is not performed. Detailed information, not precise measurements, is required. However, applications of this type must cover very limited areas because of the huge number of photographs needed for complete coverage. Befort (1986) and Ciesla (1989) provide other recent examples of this type of application. Species identification of individual trees is performed with accuracies equivalent to ground survey (e.g. Befort, 1988; Needham and Smith, 1987; Paine and McCadden, 1988). The second direction requires precise knowledge of scale (or altitude) for the purpose of making inventory measurements. These important techniques are discussed below in the forest mensuration section.

Small format photography has also proven useful and economical when used at smaller scales. Meyer (1982) reviews the utility of small format photography for many resource management activities, including forest inventory. Inexpensive small format photography can be especially useful while supplementing black and white large format photography with color and color infrared emulsions. Meyer et al. (1982) present a detailed review of utility, methods, and costs of a specific large area project using 35 mm. color infrared film. When used appropriately, the small format imagery is shown to be very economical and effective.

Small scale photography can be obtained using small format cameras or from high altitudes or space. High altitude color infrared photography acquired by either NASA or the National High Altitude Program (NHAP) has been widely applied to forest classification (e.g. Bonner, 1981; Brockhaus, 1989; Hegg et al., 1981). This small-scale, large area imagery has proven very effective for classifying large areas of forest into general forest classes. NHAP has now been replaced by NAPP--the National Aerial Photography Program, which is described by Dickson (1989). NAPP is important to forestry since it is designed to acquire periodic, high quality color infrared aerial photography of all areas of the U.S. Panoramic imagery, a special case of small scale, high altitude aerial photography, also holds promise but presents some specialized analysis and handling requirements (Befort et al., 1980; Scott et al., 1987).

Digital satellite imagery and image processing has also been applied to forest classification, including stratification for inventory purposes. Numerous forest classification projects have experimented with Landsat Multispectral Scanner (MSS) imagery (see Heller and Ulliman, 1983). The results have varied widely depending on the techniques used and the location and extent of the study regions. Heller and Ulliman (1983) and Williams and Nelson (1986) report that MSS data can be used to distinguish forest and other level I (Anderson et al., 1976) land cover classes with accuracies near 90 percent. Also, level II forest classes (essentially hardwood vs. softwood) can generally be classified with accuracies ranging from about 70 percent to 85 percent. Specification of level III forest type classes (species groupings) cannot be performed consistently. These performance levels are sufficient only for forest stratification purposes and regional projects.

Forest Mensuration

Variables Related to Volume. Forest mensuration using remote sensing is almost exclusively performed using photogrammetric measurements from aerial photographs, although some research has shown relationships between digital image values and biomass or broad volume classes (e.g. Sader, 1987). Timber volume estimation is the usual goal of inventory by remote sensing, but other purposes such as biomass or potential fuel are also possible (Meeuwig et al., 1979). Several components of timber volume may be measured from aerial photographs. These components are tree (or stand) height, crown diameter, crown closure, crown area, and number of trees. Height is the component which is best related to volume but at least one additional measurement is usually desired during volume estimation. Generally, there are two approaches to forest mensuration using aerial photography: the classic approach using large or medium scale photographs to indirectly estimate volumes or more recent techniques using very large scale photographs to produce more detailed inventory information.

Classic Approach. The classic approach is introduced by Spurr (1948), reviewed by Heller and Ulliman (1983), and presented in detail by Paine (1981). Measurements made on aerial photographs are related to volume measurements made on the ground by regression analysis techniques. The regression equations are then summarized in photo volume tables which may be developed for individual species or species groups. Two types of volume tables are used: single-tree volume tables or stand volume tables. The single-tree tables are used infrequently since accurate counts of trees are difficult to make except in open stands or with very large scale photos. Single-tree techniques also consume large amounts of time. The stand volume tables are more popular since crown density (or some estimate of stand density) is used instead of tree counts. When adjusted by ground double sampling, stand volume tables provide net results which are reliable. However, the ground sampling is expensive and makes the technique economical only for large areas. Also, the stand volume tables provide no information about the distribution of volume between species or size classes (i.e. there are no stand or stock tables). Aerial photo volume tables are readily available to foresters in the United States (e.g. Setzer and Mead, 1985) and have been proven reliable and very cost-effective. However, only limited use has been made of the techniques. This limited use has been attributed to lack of statistical control and lack of expertise with photographic measurements (Paine, 1981; Smith, 1986; Smith et al., 1986; Winterberger and Larson, 1988). Refinement and improvement of the techniques continues and application to forest inventory will increase.

Controlled-Scale Photography. Very large scale aerial photographs can provide accurate measurements of tree and stand size and volume. These techniques overcome many of the limitations of photo-based stand volume tables. The techniques have been developed mostly in Canada, based upon work reported by Sayn-Wittgenstein and Aldred (1967) and Aldred and Sayn-Wittgenstein (1972). This type of use is often at very large scales (1:500 to 1:1500) and is termed controlled-scale photography in this paper. Spencer and Hall (1988) and Befort (1988) present reviews of the development and use of controlled-scale aerial photography for forest inventory. Controlled-scale photographs can now be acquired with high quality due to improvements in film characteristics, cameras, altimeters, and acquisition techniques. These capabilities have developed over a period of about 30 years with recent significant improvements resulting from developments in altimetry, photogrammetry, and positioning technology. Controlled scale photographs have been shown to provide certain types of timber measurements (e.g. height, crown diameters, areas) with accuracies meeting or exceeding ground measurements. When used for large inventory projects, controlled-scale photographs can reduce costs by 50% even when considering the substantial investment in equipment (Spencer and Hall, 1988). Gagnon et al. (1987) report several photogrammetric procedures which reduce some of the system costs. Large scale and controlled-scale photographs have also been shown very useful while performing regeneration inventories (Kirby, 1980). Controlled-scale photography is used extensively in Canada for forest inventory but is virtually unknown in the U.S. where timber is more accessible. A recent paper (Paine and McCadden, 1988) reports the use of very large scale photographs for making detailed timber volume measurements without using scale or altitude information. This new technique appears promising for areas where tariff volume tables are used.

Inventory Sample Designs

Under most circumstances, remote sensing is unsuitable as the only source of information and measurements. However, remote sensing techniques (ranging from controlled-scale photography to digital image processing of satellite data) are desirable when used appropriately within a well-designed sampling scheme. When designed correctly, sample schemes which include remote sensing will either be less expensive for the same level of precision or more precise for the same level of expense (Thomas and DeGloria, 1979). The advantages of remote sensing become more pronounced as the region to be inventoried gets larger or more remote or both. The development of forest sampling using remote sensing continues to be refined. For example, Gering and McTague (1988) report on a technique for sampling stand density by probability proportional to size.

Appropriate sample designs are generally included in two groups: double sampling or multi-stage sampling. Double sampling refers to ground measurement of a subset of the photo plots in order to remove photo interpretation or measurement bias (Paine, 1981). Depending on the scale of the photography and the mensuration method used, double sampling can range from minimal to extensive. The degree of double sampling directly impacts the economy of using remote sensing, especially for limited areas which are accessible.

Remote sensing, including space imagery, has proven effective when included in multi-stage sampling schemes. Multi-stage techniques stratify the resource for more efficient and accurate sampling (Titus, 1979). The resource strata progress from general classes to more specific inventory categories. Remote sensing is particularly effective in grouping the resource into appropriate strata. The most general levels are derived from small scale imagery, including satellite data. The most detailed strata are usually measured on the ground or using controlled-scale photography. Multi-stage sampling for forest inventory has been applied successfully numerous times (Dempster and Scott, 1979; Hall and Hales, 1979; Harding and Scott, 1978; Hegyi, 1979; Kirby, 1980; Langley, 1969; Nichols et al., 1976; Scott et al., 1987; Winterberger and Larson, 1985).

In 1975, Titus reported a random stratified sampling experiment, using Landsat-1 multispectral scanner (MSS) digital image data to select primary sampling units, and specially acquired large-scale aerial photography over the selected sites to provide a second stage. The third sampling stage was to visit the selected plots and select sample trees. The fourth stage was to subsample these trees for precise volume measurement. The estimated sampling error was 7.8 percent, and the cost was 65 percent less than a survey of comparable precision based only on ground measurements. However, not all subsequent surveys conducted in this fashion have been so successful.

CURRENT DEVELOPMENTS IN FOREST INVENTORY

Remote sensing, photogrammetry, and other aspects of the mapping sciences are advancing rapidly. This advancement impacts all applications, including forest inventory. This section will briefly identify some of the promising current developments which will impact forest inventory.

Digital Image Acquisition and Analysis

Digital satellite imagery has now developed beyond the capabilities of Landsat MSS. The same development is true for the digital processing systems which operate on the satellite data. The most significant satellite advances are the Landsat Thematic Mapper (TM) satellite remote sensing instrument and the SPOT (Systeme Pour l'Observation de la Terre) remote sensing satellite. Improvements in capability have also occurred in the techniques and equipment used to process digital imagery, co-incident with dramatic decreases in price and difficulty of use.

The higher quality and more modern image data from the Landsat Thematic Mapper (TM) now appears to provide greater forest classification information, including level II forest land cover categories (Hopkins et al., 1988; MacKenzie, 1988). Under special conditions, species information may even be possible. However, the data are new and much more application and evaluation must be performed before definitive forest classification results can generally be attributed to TM imagery.

SPOT imagery, from an operational satellite developed by France, has generated great interest because of its impressive spatial resolution. The detail of this imagery (10 or 20 meter ground resolution, depending on imaging mode) is recognized as an important advantage for visual interpretation of forest areas (Lillesand et al., 1985). However, SPOT lacks the spectral information content of other image types for image interpretation (Chavez and Bowell, 1988; DeGloria and Benson, 1987). The combination of SPOT imagery and its high spatial resolution with TM imagery holds great promise (Chavez and Bowell, 1988).

A number of additional earth-sensing satellites have just been launched or are being manufactured for launch in the next few years. These new data sources will continue to enhance the use of satellite imagery for forest inventory. Lauer (1985) provides a brief review of these other satellites.

Image processing techniques have also evolved to improve the results which were generally attributed to MSS. With the improved image sources of today and the near future, these newer analysis procedures show great promise for detailed information. Some of the techniques relate purely to the processing of image data (e.g. Lozano-Garcia and Hoffer, 1985; Skidmore and Turner, 1988), some relate to the combination of different types of imagery (e.g. Chavez and Bowell, 1988; Paris and Kwong, 1988), and many to the combination of image data with other spatial information through the use of Geographic Information Systems (GISs--more below). All of these improvements to image processing algorithms and procedures come at a time when image processing systems are much less expensive, more available, and easier to use. Several powerful image processing systems are now available on common microcomputer and workstation platforms, often in combination with GIS capabilities.

Video Imagery

Video imagery is emerging as an inexpensive and flexible source of data for many forestry activities, including classification and monitoring. Through these and other appropriate uses, video has potential in forest inventory. Meisner (1985) and Vlcek (1983) describe the advantages, disadvantages, and use of video, including the new color infrared capability. Applications and further development with video imagery are described by Hame and Rantasuo (1988) and Yuan et al. (1987).

Radar Imagery

In recent years, imaging Radar, especially from space, has been the subject of much research related to resource classification, inventory, and monitoring. No longer is this image type considered suitable only for areas of persistent cloud cover. A wide range of research and development with Synthetic Aperture Radar (SAR) imagery for forestry purposes is evident over the past few years (Durden et al., 1989; Hoffer and Lozano-Garcia, 1987; Paris and Kwong, 1988; Richards et al., 1987; Sader, 1987; Sun and Simonett, 1988; Wu, 1987). This research indicates great utility of SAR for forestry purposes, including direct relationships between certain Radar data types and woody biomass, forest canopy structure, and forest classes. SAR is especially useful when used with other image types. Although Radar imagery is not readily available to forest managers, repetitive and detailed SAR coverage of earth from a Canadian satellite is imminent.

Analytical Photogrammetry

Volume measurements have been made using photogrammetric plotters and novel sampling schemes (Smith, 1969). However, photogrammetry is rapidly evolving in techniques and equipment related to the use of mathematics and computers and termed analytical photogrammetry. Wolf (1980) defines analytical photogrammetry as "photogrammetry in which solutions are obtained by mathematical methods." Analytical plotters have been developed to combine photogrammetric plotters and computers which perform the solutions. These plotters have only been commercially available for about 20 years and represent substantial improvements in accuracy, productivity, and flexibility. However, analytical plotters have been too expensive and complicated for forest mensuration work until very recently. Now, microcomputer-based analytical plotters are available at much lower prices while still allowing measurements with enough accuracy for forest mensuration. Warner (1988) discusses the use of low-cost analytical plotters in the forest enterprise. Analytical instruments have also been used with the controlled-scale photography projects discussed above. Other indications of the use of analytical photogrammetry for forest measurements are provided by the purchase of analytical plotters by the U.S. Forest Service (personal communication, 1989, Clearwater National Forest, Orofino, ID) and articles like the paper by Reece et al. (1983).

Laser Altimetry

The use of laser altimeters in forest inventory has already been discussed above in relation to controlled-scale photography. However, recent research has also investigated use of these instruments without photography (Baker et al., 1985; Schreier et al., 1983). The principle involves differences measured by the laser altimeters between the forest canopy and the forest floor. Tree heights, spacing, and diameters may be measured with the instruments and strip sampling techniques used to estimate volumes.

Geographic Information Systems

Geographic Information Systems (GISs) have seen accelerated use in the past few years due to improving computer technology which is also becoming less expensive. The application of GIS to forestry is too broad a subject for this paper, but mention needs to be made of the relationship between GIS and remote sensing. Digital image analysis for forest classification has consistently shown improvement when spatial data in the GIS were used during the classification (Cibula and Nyquist, 1987; Leprieur et al., 1988; Marble et al., 1983; Strahler et al., 1979). Topographic variables are consistently found to influence the accuracy of the analyses. Forest productivity has also been the subject of GIS and remote sensing work (Cook and Iverson, 1987; Craul, 1989).

Spatial Positioning

Technology for spatial location has exploded. Even though applications to forest inventory remain unreported, failure to mention these current technologies in this paper would be short-sighted.

Radio navigation (LORAN), analytical aerotriangulation, inertial surveying systems, and satellite positioning (the Global Positioning System--GPS) are the principal areas of potential. The recent refinement of GPS technology continues to reduce the cost of the equipment, the weight of the equipment and the ease with which persons can be trained to operate the equipment productively. The use of GPS for forest-related measurements is particularly advantageous since the use of intervisible marks is not necessary and the measurements can be made both day and night in all types of weather. GPS relative positioning accuracies of one part per million (ppm) or better are possible over distances from a few hundred meters to thousands of kilometers (King et al., 1985). The use of GPS in Forest Inventory will first be to provide ground control for various activities. Next, it will be used to provide airborne high-accuracy (10 cm) positioning systems to support photogrammetric mapping (Merrell et al., 1989). Its ultimate use, when cost and weight requirements are reduced to a point, will be that of a real-time mapping tool which is used to update GIS systems with map and attribute data concerning specific management parcels. As mentioned often during inventory work (e.g. Spencer and Hall, 1988), precise knowledge of plot location is desirable for revisit purposes. With these new positioning technologies, knowledge of location, especially of aircraft and photo plots, will be simply and rapidly determined.

SOME THOUGHTS ABOUT ANALOG AND DIGITAL IMAGE ACQUISITION AND ANALYSIS

The acquisition of photographic imagery of forests and other ground features is a well established technique. The data is collected using a camera, which is calibrated in terms of its radiant transmittance, and in terms of its ability to record shapes, textures, boundaries and spatially varying ground features as variations in photographic density and (in the case of color or color infrared emulsion) as variations in hue, saturation and intensity. The smallest object that may be resolved depends upon the ability of the optics and of the photographic emulsion to respond to higher spatial frequencies. All spatial variations in intensity of the upwelling radiance field may be represented by a Fourier series, whose coefficients represent amplitude (the square root of intensity) and the argument of whose sinusoidal terms relate to spatial frequency. In the limit, processed photographic emulsion records radiant intensity variations as a two-dimensional pattern of silver grains or of sensitized organic dye molecules. There is therefore an upper limit to the highest spatial frequencies that a photographic emulsion can record, determined by how coarse, or how fine the deposited photosensitive silver halide grains or organic dye molecules may be. The image is therefore a sample distribution of what is on the ground. If it were not so, there would be no limit to the number of diameters to which you could enlarge a photographic image. The ability of the optical system to transmit spatial frequencies, and of the film to record them is given by the modulation transfer function (MTF). A camera which could record the finest detail with total clarity would transmit without frequency-dependent attenuation and record all spatial frequencies in the upwelling radiance field. This would correspond to a modulation transfer function of 1.0 at all spatial frequencies. A real system has a modulation transfer function which falls off with increasing spatial frequency. This causes a loss of definition. Edges do not appear sharp, and there is a limit to how small objects may be and still be resolved. High quality films and camera systems have a modulation transfer function which falls off slowly with increasing spatial frequency.

The human interpreter uses the eye/brain system to extract meaning from an image, or to a temporal sequence of images by seeking cues from shape, size, texture, tone/color, relative position and change with time. Experience is a necessary component of the analytical process and photointerpretation is an art insofar as the sequence of decision-making procedures used by the interpreter are not clearly understood. All of us use this technique to some extent, since spatially varying information can only be presented and understood in terms of a pictorial product. Indeed, three channels of information may be superimposed and viewed simultaneously, by coding them as red, green and blue, as in the case of a color, or color infrared film.

In the case of a digital image, however, the resolution of the sensor is determined by its instantaneous field of view (IFOV) and of its ground projection (the ground instantaneous field of view: GIFOV). The radiance recorded from each sample (each GIFOV) is represented on the resulting image as a pixel. Due to oversampling, adjacent pixels will contain radiance from the same area on the ground, due to some overlap of GIFOVs. For example, although the Landsat multispectral scanner MSS had a 79m by 79m GIFOV, the sensor sampled at centers 56m apart along each scan line. The correct aspect ratio is recorded on the image by using pixels which have side dimensions in the ratio 56:79. The digital image may be reproduced at any scale until the dimensions of the pixels introduce difficulty. The radiometric fidelity of digital image data is generally better than for film. A digital imaging sensor may quantize radiance to 4096 levels, corresponding to a twelve-bit byte to describe the radiant intensity from the GIFOV in each single bandpass image under consideration. This contrasts with about a four percent precision in radiometric intensity from a photographic film.

The analysis of digital image data may be fully or partially automated. In the former case, the image data, which is generally simultaneously collected in several wavebands, yielding a multiband image, is subjected to cluster analysis, in which the data cluster into classes corresponding to like recorded spectral radiance signatures. The axes of the hyperspace consist of recorded radiance values in each of the bandpasses in which images are obtained. Alternatively, ground areas of interest, having radiance values which are discriminable are used as "training areas", and decision rules are used to decide which of the pixels of which the image consists correspond to each type of ground cover. Such techniques are referred to as feature extraction. Another set of procedures is referred to as image enhancement. Here, contrast ranges may be selected and made to fill the dynamic range of the interactive computer screen; in this way an image may be made more useful to the interpreter. Both procedures usually ultimately result in the production of a photographic (analog) image which contains that information which the interpreter considered to be useful. This information may be used as the input to a GIS. Indeed, many researchers consider that the full potential of geographic information systems and remote sensing technology can best be realized if the two technologies are integrated (Shelton and Estes, 1981).

The comparative characteristics of analog photographic images and digital images are best represented in Table 1 (Duggin and Robinove 1989).

RELATIONSHIP BETWEEN THE IMAGE AND THE GROUND

The use of ground data is necessary in stratified sampling experiments, such as volume estimation, and to train the computer in the case of supervised multichannel digital image classification. Ground data is also necessary to verify the accuracy of such classification. However, since the ground data is assumed to be completely accurate, care must be taken to ensure that the ground information is both relevant and accurate. For example, aerial photo volume tables are constructed based on ground measurements of plots on which photo measurements are then made. The photo measurements are then related to the ground values (usually by regression). If the ground data is old or contains significant measurement error, then the photo volume tables will be inaccurate for reasons beyond the use of aerial photographs.

In another example, the ground data, relating to ground features which are to be mapped on the basis of recorded multichannel radiometric differences, should be obtained at the same time as the remotely sensed image unless the features to be mapped are known to be insensitive to time and to (e.g.) weather conditions which might prevail at the time of acquisition of the two data sources. Failure to ensure this synchronization in image and in ground data acquisition might result in the comparison of two unrelated data sets. Failure to collect ground data and image data at the same time could result in the use of inappropriate ground information to perform the supervised image classification and to

Table 1. Comparative Characteristics of Analog Photographic Images and Digital Images

Photographic Image	Digital Image
Scale of the image depends upon focal length and flying height.	Scale is based solely upon dimensions of the ground resolution cell represented by each pixel. An image of any dimensions, and therefore scale, may be produced.
Enlargement is limited by the grain size of the photographic emulsion.	Enlargement is limited by a loss of detail when individual pixels distort contrast boundaries.
Radiometric calibration is difficult and depends upon accurate sensitometric and radiometric measurements. The characteristic is never linear over the entire dynamic range of the film.	Radiometric calibration may be performed, and is usually within 1% linear over the entire dynamic range of the detectors.
Vertical aerial or space photography may be rectified using photogrammetric techniques.	Digital image data may be rectified using ground control points, resampling techniques and earth rotation corrections in the case of satellite image data.

check its accuracy, with a resulting reduction in accuracy and a consequent loss of utility of the results of the image acquisition and analysis task.

An attribute of interest such as forest type or the existence of stress, is assumed to have a strong correlation with those optical properties of the ground cover which control the upwelling radiance field. This correlation is assumed constant across an image. However, there will be both random and systematic variations in this correlation across the image as pointed out by Duggin and Robinove (1989). In part, this will be due to systematic atmospheric variations with view angle (air mass) and atmospheric fluctuations. Also, there is a dependence of the reflectance and emissivity characteristics of scene elements on view and illumination angles (e.g. Coulson 1966, Kimes 1983, 1984, Kimes and Kirchner 1983, Verhoef and Bunnik 1976, 1981). This will give rise to variation in scene radiance in a systematic manner, for a horizontal scene consisting of one component. For a uniform scene composition, such as a homogeneous forest consisting of one species, the variation in radiance across a scene with view angle would be systematic. However, topography and heterogeneity in the composition and spatial distribution of scene elements, each with their own optical anisotropy, will give rise to a random variation in the upwelling radiance field. This effect will be exacerbated by varying topographic slope and aspect effects.

In order to estimate the impact of systematic and random variations across an image, one approach is to identify areas on the image which are known from other means to exclusively relate to the feature of interest. There should be several such areas distributed across the image and each area should be large enough to contain several pixels so as to avoid edge and mixed pixel effects. The variation in radiance within and between the test areas will permit estimates to be made of systematic and random variations in radiance from these test areas.

A LOOK TO THE FUTURE: IMAGING SPECTROMETRY

The selection of those optical bandpasses in which images of forest areas are simultaneously obtained are dictated by the spectral sensitivity of the monochromatic, color or false color films which are selected in the case of aerial or space photography. This restriction does not necessarily apply in the case of digital image devices such as the Landsat Thematic Mapper (TM).

The development of the High Resolution Imaging and Moderate Resolution Imaging Spectrometers (HIRIS and MODIS, respectively) to be flown on satellites in the mid-1990s promises the mapping of selected attributes providing that sufficient ground information on the optical properties related to those attributes to be mapped is available. Such ground information will permit the selection of those bandpasses most strongly correlated with the attributes to be mapped, but least strongly correlated with each other such that a resultant image may be formed from a linear, multivariate combination of the selected images. Such a resulting image should permit mapping of the extent and quantification of the attribute of interest. HIRIS will be capable of simultaneously imaging the ground with a 30m resolution from a polar orbit in 192 spectral channels between 0.4 and 2.5 micrometers. MODIS will be limited to 64 spectral channels, using a 500m spatial resolution. The data handling requirements of these new systems will be extremely challenging (Kahn and Leidecker, 1989). In both cases, it will be necessary to be selective in combining images to form a vegetation index image, on the basis of which some attributes of the forested area may be mapped. Work has been performed by Rock (1982, 1985), by Rock, Hoshizaki and Miller (1988) and by Singhroy (1989) on the effect of stress on forest canopy. Such studies are necessary for informed selection of the most appropriate spectral bandpasses for mapping a given forest attribute (e.g., Huck, et al., 1984). Cox (1983) has collated information on broad spectral bandpasses of interest in mapping vegetation attributes which are shown in Table 2.

CONCLUSION

In the past, manual image interpretation procedures involving aerial photography have been used in forest type mapping and in stratified sampling procedures for estimating timber volume.

These procedures remain valid though they are used sparingly. Recent and current developments promise to increase the efficiency, accuracy, and precision of forest inventory by using current remote sensing techniques. However, managers and inventory specialists must be aware of the techniques and the proper procedures for implementing them. Unfortunately, remote sensing is still considered a novel, complicated, and unproven tool for forestry applications. This blanket reputation is undeserved and should be replaced by cautious acceptance of techniques which can improve the amount of reliable information available to the modern resources manager and inventory specialist. The key requirement is reliable information. The methods surveyed in this paper will contribute better, more, and less expensive data to resources management.

Table 2. Spectral Regions for Remote Sensing of Vegetation*

Wavelength (m)	Type of Feature	Value
0.440-0.500	Absorptance	Detection of changes in chlorophyll/carotenoid ratios (related to stress)
0.650-0.700	Absorptance	Detection of chlorophyll states as well as tannin and anthocyanin content. Initial stress detection.
0.700-0.750	Reflectance	Senescence detection. Detection of dead or dormant vegetation.
0.800-0.840	Absorptance	Possibly related to leaf anatomy and/or state of hydration.
0.865	Reflectance	Height of feature may be useful in species discrimination.
0.940-0.980	Absorptance	Shifts in this minor water absorption band may be useful in species discrimination and determination of hydration state.
1.060-1.100	Reflectance	Shifts in peaks may be related to leaf anatomy and/or morphology. May be useful for species discrimination
1.140-1.220	Absorptance	Shifts in this minor water absorption band may be useful in species discrimination and determination of hydration state.
1.250-1.290	Reflectance	Height of this feature very useful for species discrimination of senescent forest species. A ratio of this feature with the one at 1.645 offers a good indication of moisture content and thus stress.

*Cox, 1983

1.630-1.660	Reflectance	An indication of moisture content of leaf. May also be an indicator of variation in leaf anatomy. May be useful for species discrimination. An indicator of leaf moisture content when used as a ratio with the 1.270 data above.
2.190-2.300	Reflectance	An indicator of moisture content. May also be of value in species discrimination.
3.000-5.000 8.000-14.000	Emittance	Little is known concerning the optimal thermal wavelengths for studying different vegetation parameters. This is an area that needs further study.

ACKNOWLEDGEMENTS

The authors appreciate the assistance of the following people in providing information and material which was used in the preparation and presentation of this paper:

William A. Kennedy, Earth Observation Satellite Company
Jeff Liedtke, International Imaging Systems, Inc.
Locke Stewart, Goddard Space Flight Center
Mildred Summerville, Goddard Space Flight Center

REFERENCES

Aldrich, R.C., 1979: Remote Sensing of Wildland Resources: A State-of-the-Art Review. General Technical Report RM-71, Rocky Mountain Forest and Range Experiment Station, USDA Forest Service, Fort Collins, CO, 80526, 56 pp.

Aldrich. R.C., R.A. Harding, and D.P. Paine, 1984: Remote Sensing. Section 20 of The Forestry Handbook, 2nd Edition, K.F. Wenger, ed., John Wiley & Sons, New York, pp. 1117-1163.

Anderson, J.R., Hardy, E.E., Roach, J.T. and Witmer, R.E., 1976: A Land Use and Land Cover Classification System for Use with Remote Sensor Data. Geological Survey Professional Paper 964, 28 pp.

Avery, T.E., 1978: Forester's Guide to Aerial Photo Interpretation. Agriculture Handbook 308, USDA Dept. of Ag., Forest Service, 41 pp.

Baker, R.D., LeSchack, L.A. and Whatley, M.C., 1985: The pulsed airborne laser profiler: a new system to quantify and evaluate forests. Proc. Pecora 10, Aug. 20-22, Colorado St. U., Fort Collins, CO. pp. 429-430.

Befort, W., 1988: Controlled-scale aerial sampling photography: development and implications for multiresource inventory. J. of Forestry, pp. 21-28.

Befort, Wm., 1986: Large-scale sampling photography for forest habitat-type identification. Photogram. Eng. and Rem. Sens., Vol. 52, No. 1, pp. 101-108.

Befort, W.A., Heller, R.C., and Ulliman, J.J., 1980: Viewing and handling panoramic aerial photographs. Tech. Rept. 7, Forest, Wildlife and Range Experiment Station, Univ. of Idaho, Moscow, ID 83843, 16 pp.

Bonner, K.G., 1981: Riparian vegetation mapping in northeastern California using high altitude color infrared aerial photography. Eighth Biennial Workshop on Color Aerial Photography in the Plant Sciences and Related Fields, Luray, VA, pp. 29-37.

Brockhaus, J.A., 1989: An assessment of remotely sensed imagery for use in hardwood stand density distribution mapping in central California. Tech. Papers, 1989 ASPRS/ACSM Annual Convention, Apr. 2-7, Baltimore, Vol. 3, pp. 109-117.

Burkhart, H.E., J.P. Barrett, H.G. Lund, 1984: Timber Inventory. Section 7 of the Forestry Handbook, 2nd Edition, K.F. Wenger, ed., John Wiley & Sons, New York, pp. 361-412.

Chavez, P.S. Jr., and Bowell, J.A., 1988: Photogram. Eng. and Rem. Sens., Vol. 54, No. 12, pp. 1699-1708.

Ciesla, W.M., 1989: Aerial photos for assessment of forest decline--a multinational overview. J. of Forestry, pp. 37-41.

Cibula, W.G. and Nyquist, M.O., 1987: Use of topographic and climatological models in a geographical data base to improve landsat MSS classification for Olympic National Park. Photogram. Eng. and Rem. Sens., Vol. 53, No. 1, pp. 67-75.

Cook, E.A. and Iverson, L.R., 1987: Estimating forest productivity in southern Illinois using landsat thematic mapper data and geographic information system analysis techniques. Tech. Papers, ASPRS/ACSM Convention, March 29-Apr. 3, Baltimore, pp. 255-265.

Coulson, K.L., 1966: Effects of reflection properties of natural surfaces in aerial reconnaisance. Appl. Optics, 5, 905-917.

Cox, S.C., (ed), 1983: The multispectral imaging science working group. Final Report, NASA Conference, Publication 2260, 3.20.

Craul, P., 1989: personal communication. SUNY College of Env. Sci. and For., Syracuse, NY 13210.

DeGloria, S.D. and Benson, A.S., 1987: Interpretability of advanced SPOT film products for forest and agricultural survey. Photogram. Eng. and Rem. Sens., Vol. 53, No. 1, pp. 37-44.

Dempster, W.R. and Scott, C.A., 1979: A multilevel approach to acquiring resource data for timber management planning. Proc. Practical Applications of Remote Sensing to Timber Inventory, workshop, Sept. 26-28, Edmonton, Alberta, Canada, Information Report NOR-X-224, pp. 127-137.

Dickson, R.A., 1989: Products available from NAPP imagery. Tech. Papers, 1989 ASPRS/ACSM Annual Convention, Apr. 2-7, Baltimore, Vol. 1, pp. 238-242.

Duggin, M.J. and Robinove, C.J., 1989: Assumptions implicit in remote sensing data acquisition and analysis. In publication.

Durden, S.L., van Zyl, J.J. and Zebker, H.A., 1989: Modeling and observation of the radar polarization signature of forested areas. Proc. IEEE Trans. Geosci. and Rem. Sens., 27, 290-301.

Gagnon, P.A., Mainville, M., Agnard, J.P. and Tomlinson, J., 1987: Practical photogrammetric procedures for large-scale forest inventories. Tech. Papers, 1987 ASPRS/ACSM Annual Convention, March 29-Apr. 3, Baltimore, Vol. 2, pp. 8-16.

Gering, L.R. and McTague, J.P., 1988: Estimating stand density of loblolly pines in northern Louisiana using aerial photographs and probability proportional to size. Tech. Papers, 1988 ACSM/ASPRS Annual Convention, March 13-18, St. Louis, Vol. 4, pp. 222-228.

Hall, J.A. and Hales, M.T., 1979: Multistage inventory technique case study-McCain Valley. Seventh Biennial Workshop on Color Aerial Photography in the Plant Sciences and Related Fields, U.C. Davis, May 15-17, pp. 119-124.

Hame, T. and Rantasuo, M., 1988: Shuttered camera--aerial color video imaging in the visible and near infrared. Photogram. Eng. and Rem. Sens., Vol. 54, No. 12, pp. 1738-1738.

Harding, R.A. and Scott, R.B., 1978: Forest Inventory with Landsat, Phase II. Washington Forest Productivity Study, State of Washington, Dept. of Natural Resources, Olympia, WA 98504, 221 pp.

Hegg, K.M., Larson, F.R., Mead, D.R. and Winterberger, K.C., 1981: Multi-resource inventory in interior Alaska. Eighth Biennial Workshop on Color Aerial Photography in the Plant Sciences and Related Fields, Luray, VA, pp. 59-65.

Hegyi, F., 1979: A new approach to forest inventory using remote sensing. Proc. Practical Applications of Remote Sensing to Timber Inventory, workshop, Sept. 26-28, Edmonton, Alberta, Canada, Information Report NOR-X-224, pp. 110-115.

Heller, R.C. and Ulliman, J.J., 1983: Forest resource assessments, Chapter 34. The Handbook of Remote Sensing, ed. R.N. Colwell, American Society for Photogrammetry, Falls Church, VA, pp. 2229-2324.

Hoffer, R.M. and Lozano-Garcia, D.F., 1987: Definition of forest stand characteristics based on multi-incidence angle SIR-B data. Tech. Papers, ASPRS/ACSM Convention, March 29-Apr. 3, Baltimore, pp. 296-304.

Hopkins, P.F., Maclean, A.L. and Lillesand, T.M., 1988: Assessment of thematic mapper imagery for forestry applications under lake states conditions. Photogramm. Eng. and Rem. Sens., Vol. 54, No. 1, pp. 61-68.

Huck, F.O., Davis, R.E., Fales, C.L., Aherron, R.M., Arduini, R.F. and Samms, R.W., 1984: Study of remote sensor spectral responses and data processing algorithms for feature classification. Opt. Eng., 23, 650.

Husch, B., Miller, C.I., Beers, T.W., 1972: Forest Mensuration, The Ronald Press Co., New York, NY, 410 pp.

Kahn, R. and Leidecker, H., 1989: The crush of new earth science data knocking at our door. Renewable Resources J., Vol. 7, No. 1, pp. 8-13.

Kimes, D.S., 1984: Modeling the directional reflectance from complete homogeneous vegetation canopies with various leaf orientation distributions. J. Opt. Soc. Am., 74, 725-737.

Kimes, D.S., 1983: Dynamics of directional reflectance factor distributions for vegetation canopies. Appl. Optics, 22, 1364-1372.

Kimes, D.S. and Kirchner, J.A., 1983: Directional radiometric measurements of row crop temperatures. Int. J. Remote Sensing, 4, 299.

King, R.W., Masters, E.G., Rizos, C., Stolz, A., and Collins, J., 1985: Surveying with GPS. The University of New South Wales, Kensington, NSW, Australia, 2033.

Kirby, C.L., 1980: A camera and interpretation system for assessment of forest regeneration, Information Report NOR-X-221, Northern Forest Research Centre, Edmonton, Alberta, 8 p.

Langley, P.G., 1969: New multi-stage sampling techniques using space and aircraft imagery for forest inventory. Proc. Sixth Int. Symp. on Remote Sensing of Environ., Oct. 13-16, Ann Arbor, MI, Vol. II, pp. 1179-1192.

Lauer, D.T., 1985: New horizons in remote sensing for forest and range resource management. Proc. Pecora 10, Aug. 20-22, Colorado St. U., Fort Collins, CO, pp. 30-41.

Leprieur, C.E. and Durand, J.M., 1988: Influence of topography on forest reflectance using landsat thematic mapper and digital terrain data. Photogram. Eng. and Rem. Sens., Vol. 54, No. 4, pp. 491-496.

Lillesand, T.M., Hopkins, P.F., Buchheim, M.P., and Maclean, A.L., 1985: The potential impact of thematic mapper, SPOT and microprocessor technology on forest type mapping under lake states conditions. Proc. Pecora 10, Aug. 20-22, Colorado St. U., Fort Collins, CO, pp. 43-57.

Lozano-Garcia, D.F., and Hoffer, R.M., 1985: Evaluation of a layered approach for classifying multitemporal landsat MSS data. Proc. Pecora 10, Aug. 20-22, Colorado St. U., Fort Collins, CO, pp. 189-199.

MacKenzie, M., 1988: Remote sensing of the deciduous vegetation of Great Smoky Mountains National Park. Tech. Papers, 1988 ACSM/ASPRS Annual Convention, March 13-18, St. Louis, Vol. 4, pp. 212-221.

Marble, D.F., Peuquet, D.J., Boyle, A.R., Bryant, N., Calkins, H.W., Johnson, T. and Zobrist, A., 1983: Geographic Information Systems in Remote Sensing, Chapter 22. The Manual of Remote Sensing, ed. R.N. Colwell, American Soc. for Photogrammetry, Falls Church, VA, pp. 923-958.

Meeuwig, R.O., Miller, E.L. and Budy, J.D., 1979: Estimating pinyon and juniper fuel and biomass from aerial photographs. Research Note, USDA Forest Service Int-274, Intermountain Forest and Range Experiment Station, Ogden, UT 84401, 9 pp.

Meisner, D., 1985: Color infrared aerial video: applications in natural resource management. Proc. Pecora 10, Aug., 20-22, Colo. St. Univ., Fort Collins, CO, pp. 80-86.

Merrell, R., Leach, M.P., and Clynch, J.R., 1989: Development of an Operational GPS/Photogrammetric Surveying Capability. Int. Geodetic Symposium on Satellite Positioning, Physical Science Laboratory, New Mexico St. Univ., Las Cruces, NM, 88003-0002.

Meyer, M.P., 1982: Place of small-format aerial photography in resource surveys, J. of Forestry, pp. 15-17.

Meyer, M., Opseth, S., Moody, N. and Bergstrom, L., 1982: Large-area forest land management applications of small scale cir 35mm aerial photography, Final Report, Remote Sensing Laboratory, College of Forestry and the Agricultural Experiment Station, Univ. of Minn., St. Paul, MN, 23 p.

Moessner, K.E., 1953: Photo interpretation in forest inventories. Photgram. Eng., Vol. 19, No. 3, pp. 496-507.

Needham, T.D. and Smith, J.L., 1987: Stem count accuracy and species determination in loblolly pine plantations using 35-mm aerial photography. Photogram. Eng. and Rem. Sens., Vol. 53, No. 12, pp. 1675-1678.

Nichols, J.D., Harding, R.A., Scott, R.B. and Edwards, J.R., 1976: Forest inventory of western Washington by satellite multi-stage sampling. Proc. Am. Soc. of Photogrammetry, Fall Convention, Seattle, WA, Sept. 28-Oct. 1, pp. 180-217.

Paine, David P., 1981: Aerial Photography and Image Interpretation for Resource Management. John Wiley & Sons, Inc., 571 pp.

Paine, D.P. and McCadden, R.J., 1988: Simplified forest inventory using large-scale 70-mm photography and tarif tables. Photogram. Eng. and Rem. Sens., Vol. 54, No. 10, pp. 1423-1427.

Paris, J.F. and Kwong, H.H., 1988: Characterization of vegetation with combined thematic mapper (TM) and shuttle imaging radar (SIR-B) image data. Photogram. Eng. and Rem. Sens., Vol. 54, No. 8, pp. 1187-1193.

Reece, J.W., Iff, R.H. and Sirois, D.L., 1983: An inexpensive photogrammetry system for obtaining 3-D spatial coordinates in dynamic testing conditions. Forest Sci., Vol. 29, No. 2, pp. 341-350.

Richards, J.A., Sun, G.Q. and Simonett, D.S., 1987: L-band radar backscatter modeling of forest stands. IEEE Trans. Geosci. Rem. Sens., vol. GE-25, pp. 487-498.

Rock, B.N., 1985: Remote detection of geobotanical anomalies associated with hydrocarbon microseepage using thematic mapper simulator (TMS) and airborne imaging spectrometer (AIS) data. Remote Sensing for Geologic Mapping, 18, ed. by Teleki and Webber, p. 299, Int. Union of Geological Sciences.

Rock, B.N., 1982: Mapping of deciduous forest cover using simulated Landsat DTM data. Proc. Int. Symp. on Geoscience and Remote Sensing, 5, 3.1.

Rock, B.N., Hoshizaki, T. and Miller, J.R., 1988: Comparison of in situ and airborne spectral measurements of the blue shift associated with forest decline. Rem. Sens. of Environ., 24, 109-127.

Sader, S.A., 1987: Forest biomass, canopy structure, and species composition relationships with multipolarization L-band synthetic aperture radar data. Photogram. Eng. and Rem. Sens., Vol. 53, No. 2, pp. 193-202.

Sayn-Wittgenstein, L., 1978: Recognition of Tree Species on Aerial Photographs. Forest Management Institute, Ottawa, Ontario, Information Report FMR-X-118, Canadian Forestry Service, Dept. of the Environment, 97 pp.

Sayn-Wittgenstein, L., 1960: Recognition of Tree Species on Air Photographs by Crown Characteristics. Forest Research Branch, Technical Note No. 95, Canada Department of Forestry, 56 pp.

Sayn-Wittgenstein, L. and Aldred, A.H., 1972: Tree size from large-scale photos. Photogram. Eng., Vol. 38, No. 10, pp. 971-973.

Sayn-Wittgenstein, L. and Aldred, A.H., 1967: Tree volumes from large-scale photos. Photogram. Eng., Vol. 33, No. 1, pp. 69-73.

Schreier, H., Lougheed, J., Gibson, J. and Russell, J., 1983: Calibrating an airborne laser profiling system. Proc. RNRF Symp. on the Application of Remote Sensing to Resource Management, Seattle, WA, May 22-27, pp. 109-115.

Scott, C.T., Schreuder, H.T. and Griffith, D.M., 1987: A comparison of optical bar, high-altitude and black-and-white photography in land classification. Photogram. Eng. and Rem. Sens., Vol. 53, No. 2, pp. 203-206.

Setzer, T.S. and Mead, B.R., 1985: Aerial photo stand volume tables for southeast Alaska. Proc. Pecora 10, Aug. 20-22, Colorado St. U., Fort Collins, CO. pp. 407-408.

Shelton, R.C. and Estes, J.E., 1981: Remote sensing and geographic information systems: an unrealized potential. Geo-Processing, 1, 395-420.

Singhroy, V.H., 1989: Ph.D. Thesis. College of Environmental Science and Forestry, S.U.N.Y., Syracuse, NY 13210.

Skidmore, A.K. and Turner, B.J., 1988: Forest mapping accuracies are improved using a supervised nonparametric classifier with SPOT data. Photogram. Eng. and Rem. Sens., Vol. 54, No. 10, pp. 1415-1421.

Smith, D.U., 1969: Timber volume with a Kelsh plotter. Photogramm. Eng., 35, 363-365.

Smith, J.L., 1986: Evaluation of the effects of photo measurement errors on predictions of stand volume from aerial photography. Photogram. Eng. and Rem. Sens., Vol. 52, No. 3, pp. 401-410.

Smith, J., Bleier, M. and Oderwald, R., 1986: The data dilemma: how to properly construct and utilize aerial photo volume tables. Tech. Papers, 1986 ACSM/ASPRS Annual Convention, March 16-21, Washington, D.C., Vol. 4, pp. 286-293.

Spencer, Ray D. and Hall, R.J., 1988: Canadian large-scale aerial photographic systems (LSP). Photogram. Eng. and Rem. Sens., Vol. 54, No. 4, pp. 475-482.

Spurr, Stephen H., 1948: Aerial Photographs in Forestry. The Ronald Press Company, New York, 340 pp.

Strahler, A.H., Logan, T.L. and Woodcock, C.E., 1979: Forest classification and inventory system using Landsat, digital terrain, and ground sample data. Proc. Thirteenth Int. Symp. on Rem. Sens. of Environ., Ann Arbor, MI, pp. 1541-1557.

Sun, G. and Simonett, D.S., 1988: A composite L-band HH radar backscattering model for coniferous forest stands. Photogram. Eng. and Rem. Sens., Vol. 54, No. 8, pp. 1195-1201.

Talmadge, D.A. and Curran, P.J., 1986: Remote sensing using polarized light. Int. Jol. of Rem. Sens., 7, 47-64, 1986.

Thomas, R.W. and DeGloria, S.D., 1979: Large-scale aerial photography as a sample stage in wildland resource inventories. Proc. of the Seventh Biennial Workshop on Color Aerial Photography in the Plant Sciences, May 15-17, 1979, Univ. of Calif., Davis, CA, pp. 15-33.

Titus, S.J., 1979: Multistage sampling: what's it all about? Proc. Practical Applications of Remote Sensing to Timber Inventory, workshop, Sept. 26-28, Edmonton, Alberta, Canada, Information Report NOR-X-224, pp. 116-123.

Tueller, P.T., 1979: Some aspects of the use of dichotomous keys to aid in the interpretation of color aerial photographs for vegetation mapping. Seventh Biennial Workshop on Color Aerial Photography in the Plant Sciences and Related Fields, U.C. Davis, May 15-17, pp. 189-200.

Vanderbilt, V.C., Grant, L. and Daughtry, C.S.T., 1985: Polarization of light scattered by vegetation. Proc. IEEE, 73, 1012-1024.

Verhoef, W. and Bunnik, N.J.J., 1981: Influence of crop geometry on multispectral reflectance determined by the use of canopy reflectance models. Proc. of the Int. Colloquium on Signatures of Remotely Sensed Objects, pp. 273-290.

Verhoef, W. and Bunnik, N.J.J., 1976: The spectral directional reflectance of row crops. Part 1: consequences of non-Lambertian behaviour for automatic classification. Part 2: measurements on wheat and simulations by means of a reflectance model for row crops. NIWARS Tech. Rept., Delft, The Netherlands; NIWARS-Publ. 35.

Vlcek, J., 1983: Videography: some remote sensing applications. Tech. Papers, 49th Annual Meeting, ACSM/ASP Convention, March 13-18, Washington, D.C., pp. 63-69.

Warner, W.S., 1988: A PC-based analytical plotter designed for forestry photogrammetry. J. of Forestry, Vol. 86, No. 12, pp. 36-39.

Williams, D.L. and Nelson, R.F., 1986: Use of remotely sensed data for assessing forest stand conditions in the eastern United States. IEEE Transactions on Geosci. and Remote Sensing, Vol. GE-24, No. 1, pp. 130-138.

Winterberger, K.C. and Larson, F.R., 1988: Measuring crown cover in interior Alaska vegetation types. Photogram. Eng. and Rem. Sens., Vol. 54, No. 3, pp. 385-387.

Winterberger, K.C. and Larson, F.R., 1985: Multiphase sample of vegetation and related resources using landsat, aerial photo, and ground survey data. Proc. Pecora 10, Aug. 20-22, Colorado St. U., Fort Collins, CO, pp. 404-405.

Wolf, P.R. and others, 1980: Definitions of terms and symbols used in photogrammetry, Chapter XIX. Manual of Photogrammetry, C.C. Slama, ed., Am. Soc. for Photogrammetry and Remote Sensing, Falls Church, VA, 1056 pp.

Wu, S.T., 1987: Parametric analysis of synthetic aperture radar data for characterization of deciduous forest stands. Tech. Papers, ASPRS/ACSM Convention, March 29-Apr. 3, Baltimore, pp. 320-328.

Yuan, X., King, D., Cadeau, F. and Vlcek, J., 1987: Multispectral video survey of a northern Ontario forest. Tech. Papers, ASPRS/ACSM Convention, March 29-Apr. 3, Baltimore, pp. 140-145.

The Use of Aerial Photographs and Angle-Gauge Sampling of Tree Crown Diameters for Forestry Inventory

Lawrence R. Gering
Assistant Professor of Forest Biometrics
School of Forestry
Louisiana Tech University
Ruston, LA 71272

Dennis M. May
Kurt B. Teuber
Research Foresters
Southern Forest Experiment Station
USDA Forest Service
Starkville, MS 39759

ABSTRACT

The Forest Inventory and Analysis unit of the Southern Forest Experiment Station is charged with conducting continuous inventories of the forest resources of the Midsouth. Techniques that offer innovative approaches for improving the efficiency of these inventories are in demand. One new approach for estimating the density of forest stands involves the derivation of a technique for obtaining stand density from aerial photographs based on the principles of selection with probability proportional to size. This is accomplished by the development and use of an aerial-photo angle-gauge that is used in a procedure very similar to the ground point-sampling technique proposed by Bitterlich (1948).

Various aspects of this method have been investigated by McTague (1988) and Gering and McTague (1988); they concluded that further research is needed to compare this method of forest inventory with the traditional ground-based methods for areas in the South. The purpose of the study reported here was to compare the results of an inventory based on the angle-gauge method with those of a ground-based inventory conducted by the Forest Inventory and Analysis unit for one county in the Midsouth.

INTRODUCTION

Continuous forest inventories of the Midsouth are conducted by the Forest Inventory and Analysis unit of the Southern Forest Experiment Station. Techniques that offer new and innovative approaches for improving the efficiency and accuracy of these inventories are in demand. Procedures based on remotely sensed data may fulfill this need by improving the timeliness of current ground-based forest surveys and by allowing efficient data collection in geographically remote locations (such as west Texas and Oklahoma).

The ability to obtain reliable measurements of forest stand characteristics from aerial photographs has long been recognized by both researchers and practicing foresters. However, the use of remotely sensed data often has been limited to providing descriptions of land cover in the form of maps and summary statistics. Smith (1986) noted that the concept of estimating quantitative forest stand characteristics from aerial photographs has not yet reached its fullest potential.

One approach for estimating the density of forest stands was recently introduced by McTague (1988). It involved the derivation of a technique for obtaining stand density from aerial photographs based on the principles of selection with probability proportional to size. Results of his study included

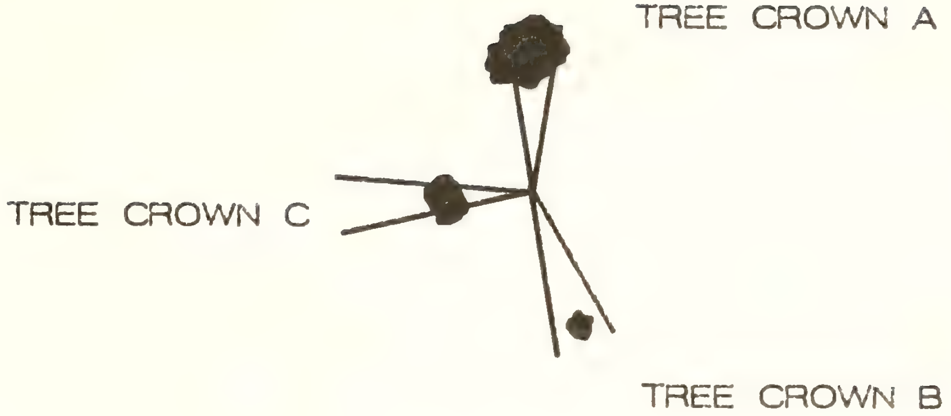


Figure 1. The crowns of trees A and C subtend an angle larger than that of the angle-gauge. Stand basal area (sq ft/acre) = (2 x BAF).

the development of an aerial-photo angle-gauge that is used in a procedure very similar to the ground point-sampling technique proposed by Bitterlich (1948) and now widely used by foresters. For the photo-based method, the angle-gauge, printed on transparent film, is rotated 360 degrees about the point center on the photograph and a count is made of all tree crowns that subtend an angle larger than that of the angle-gauge (Figure 1). McTague concluded that this method of estimating stand density is suited to the ponderosa pine (Pinus ponderosa Engelm.) type of northern Arizona.

In another study, Gering and McTague (1988) calculated the dimensions of an aerial-photo angle-gauge for estimating stand density from aerial photos of loblolly pine (Pinus taeda L.) sites in northern Louisiana. They concluded that the procedure appears to have great potential because it is quick, relatively simple and eliminates direct measurements of plot areas or tree dimensions. They also noted that more research was needed to compare this method of estimating forest inventory with the traditional ground-based methods used for areas in the Midsouth.

The objective of this study was to construct angle-gauge(s) based on the data and statistical correlations for ground-measured tree and stand variables (particularly crown width to dbh) for west-central Tennessee (Hardin County). The angle-gauge sampling technique was then used to inventory a second west-central Tennessee county (Wayne County); these results were then compared with those obtained from the ground-based inventory conducted by the U.S. Forest Service.

METHODS

Aerial photographs of west-central Tennessee (Hardin and Wayne Counties) were flown in the Fall of 1988. Color prints (with a nominal photo scale of 1:4800) were provided to crews from the U.S. Forest Service who were collecting ground-plot data during the Tennessee re-survey. Each sample point was located on the photographs and baselines were established for scale determination and photo orientation.

There were 37 plot clusters (10 sample points per cluster), with corresponding photographic coverage, established in Hardin County. All 370 sample points have been identified on the aerial photographs; this was a relatively easy task because of the uniform plot layout used by the field crews. Individual tree crowns were also identified on the photos and assigned a code number which corresponded to the tally sheets containing the ground-measured data. Crown widths were estimated using a 7x-power monoscopic comparator with a recticle illustrating circular diameters. Measurement units were converted to feet using the exact photo scale (as determined using the established baseline).

Data from Hardin County were used to develop correlations based on ground-measured dbh and photo-measured crown width. These correlations enable the construction of the angle-gauges, using established techniques, for forest stands located in west-central Tennessee.

A second west-central Tennessee county (Wayne County) has similar terrain and vegetative cover as the adjacent Hardin County. A forest inventory of Wayne County will be conducted based on aerial photographs and samples obtained using the angle-gauge(s) developed for forest stands of west-central Tennessee. By doing this, the angle-gauge(s) will not be tested against the data from which they were developed. The results of this inventory will be compared to the results of the more traditional ground-based inventory.

DATA DESCRIPTION

Previous work on angle-gauge sampling (McTague, 1988; Gering and McTague, 1988) was based on the sampling of a single tree species. This study, however, is concerned with the forest resources of an entire county and is not limited to a single tree species. Hardin County data included 748 live trees representing 5 coniferous species, predominantly loblolly pine (Pinus taeda L.), and 37 deciduous species, predominantly oak (Quercus spp.), American beech (Fagus grandifolia Ehrh.) and hickory (Carya spp.). There are also two categories of stand size class (pole and sawtimber) and three categories of crown class (dominant, codominant and intermediate).

The importance of sorting the data into similar classes prior to developing correlations becomes evident. Initial species sorting has been based on 4 general forest types: loblolly-shortleaf pine, oak-pine, oak-hickory, and oak-gum-cypress. Sorting within these forest types has also been done by stand size class and crown class. Additional categories, based on a single-species grouping such as oak or loblolly pine, have also been created by sorting the data set. As a result of the sorting, 30 separate data categories have been created and correlations between ground-measured dbh and photo-measured crown width have been developed.

The next phase of the study will be to analyze the correlations and determine whether any or all are significantly different. The goal is to minimize the required number of angle-gauges. Ideally, a single angle-gauge would be suitable for all species and classes. However, the likelihood of this result is small. Hopefully, a reasonable number (perhaps 3 or 4) of different angle-gauges will be able to adequately sample aerial photographs of west-central Tennessee.

A similar study for two parishes in northern Louisiana (Claiborne and Union Parishes) has also been established. This second study is based entirely on field-measured data (suitable photographic coverage is not available) which represent 1319 trees. Correlations for data groupings are also being developed and analyzed. A comparison of the angle-guages for Louisiana and those for Tennessee will be made to determine if significant differences are present. This will indicate whether an angle-gauge developed for a given set of conditions is valid over a broad geographic range or if it is suited only for a limited area.

CONCLUSION

Obviously, the study reported here has not yet been completed. The logistic problems of coordinating data collection by several different methods and joining the data for analysis and comparison have delayed this project. Rather than a review of partially completed work, the intent of this paper has been to describe an inventory technique, based on simple aerial photographs, that may interest both research- and field-oriented foresters. The angle-gauge sampling technique is conceptually sound and application is relatively simple. However, any decision concerning future use of this technique must consider the comparison of the angle-gauge inventory with the traditional ground-based inventory, as established in this study.

REFERENCES

Bitterlich, W. 1948. Die Winkelzahlprobe. Allgem. Forest-u, Holzw. Ztg. 59 (1/2):4-5.

Gering, Lawrence R. and McTague, John Paul. 1988. Estimating stand density of loblolly pine in northern Louisiana using aerial photographs and probability proportional to size. In: Technical Papers of 1988 Annual Convention of the American Society for Photogrammetry and Remote Sensing; Vol. 4: 222-228.

McTague, J.P. 1988. Estimation of stand density with probability proportional to size from aerial photography. West. J. Appl. For. 3(3):89-92.

Smith, J.L. 1986. Evaluation of the effects of photo measurement errors on predictions of stand volume from aerial photography. Photo. Eng. and Rem. Sens. 52:401-410.

ACKNOWLEDGEMENT

The authors would like to express their thanks to Catherine C. Taylor and Lee Reynaud, undergraduate students in the School of Forestry at Louisiana Tech University. This study was supported by as a cooperative research project between the USDA Southern Forest Experiment Station and Louisiana Tech University. Additional funding was provided by the McIntire-Stennis Cooperative Forestry Research Program under Public Law 87-788 and the State of Louisiana.

Satellite Remote Sensing and Inventory of Minnesota Forest Resources

M.E. Bauer, Professor; T.E. Burk, Associate Professor; A.R. Ek, Professor
S.C. Ahearn, Assistant Professor; and L.P. Queen, Research Associate
Department of Forest Resources
and
Remote Sensing Laboratory
University of Minnesota
St. Paul, Minnesota 55108

ABSTRACT

Forests, covering nearly a third of the land area of Minnesota, are a significant contributor to the state's economy. The synoptic, but at the same time high resolution, view of the Earth's surface provided by Landsat TM data provides a potential means to obtain statewide inventory information at reduced cost. The objective of this investigation is to develop and test procedures using multispectral satellite data together with improved classification and sampling designs to inventory and map forest resources in Minnesota. Additional objectives are to research alternative new approaches to data classification, sampling, and estimation and determine to what degree satellite data classifications can be used to obtain information on stand density, size class, age, and disturbance. The approach is to define, implement and test classification, sampling and estimation procedures in an operational setting while conducting research and development of innovative methods which can be incorporated at later stages. Evaluation of the resulting system by the Minnesota Department of Natural Resources is a key element of the project. Successful accomplishment of the project will result in operational use of satellite data to inventory forests in Minnesota, along with potential application in other states.

INTRODUCTION

Forests are a key component of Minnesota's resources and its economy. At the time of the last statewide inventory in 1977 forests covered nearly 17 million acres or about one-third of the state's total land area. The forest-based portion of the economy, including logging, lumber, pulp and paper mills, and board and furniture manufacturing, accounts for approximately 4% of the gross state output and 7% of its net exports. In addition, forests make a major contribution to the state's recreation and tourism industries. Although forests obviously are a significant contributor to the state's economy, it is also clear that Minnesota's timber resources have much more potential to contribute to the state's economic growth and development (Lewis 1984). Thus, an important goal is to identify and manage the state's renewable forest resources to encourage expansion of the forestry sector of the state economy. To achieve the full potential of these resources requires up-to-date, accurate information on which to base economic development and forest management decisions.

The state of Minnesota is actively involved in promoting industrial expansion in the forest products industry. Decisions on over $500 million of capital investment for expansion or new plant locations are contingent on current resource assessment data. However, disagreement exists over how much forest resource is available for these developments. Because of an imbalance in age class distribution (mature and over-mature classes predominate) substantial increases in harvesting could likely be sustained in the immediate future, but within the next 30 to 40 years annual removals may surpass annual growth for some species. But, others argue that as utilization standards increase and harvesting is extended to mixed species stands, more timber will become available. Also, increased harvesting rates will result in more young stands with faster growth rates than the current mature stands. Clearly, sound decisions concerning this resource base and the rates of harvesting and renewal require accurate, up-to-date information. Existing forest inventory area data are simply too old in many respects to fully address these issues.

Although forest stand growth models applied to inventory plot records have become increasingly important for updating inventory information and for projecting future forest conditions (Ek 1983), such updates are of limited value in estimating overall forest change because of their concentration on existing forested plots. Total forest area and area by cover type also change because of farmland abandonment, harvesting and urban development, etc. Such changes are extremely difficult to model; that is the major reason for the emphasis on forest area estimation in this project.

Additional rationale for the emphasis on area estimation is that conditions within strata change little over time, but the areas and locations of strata do. For example, the mean characteristics of aspen pole-sized stands of medium density (one such strata) have changed little over centuries. Further, we have much ground sample plot information to indicate the mean of the strata. Hansen and Burk (1986) have evaluated the use of prior surveys for estimating mean characteristics of strata and found high levels of consistency between surveys within cover type-size-density strata for variables such as mean volume per acre, average tree size, and percent of volume by cover type species. The major inventory problem is thus to estimate the area and location of these strata. Given that, area by cover type-size-density class (strata area) times the mean characteristics of such strata provides estimates of overall timber volumes.

Current Forest Inventory in Minnesota

The ongoing forest inventory program in Minnesota is based on data from two types of surveys: (1) a statewide assessment of all forest land for broad scale planning and information purposes and (2) surveys of individual tracts and stands for actual land management. The first type of survey (Phase I) has traditionally been conducted by the U.S. Forest Service, North Central Forest Experiment Station (Jakes 1980). The fourth such survey completed was in 1977 and involved multiphase sampling. The first phase involved aerial photo interpretation of 277,000 one acre plots to determine forest vs. non-forest area. The second phase consisted of field checking over 35,000 of these plots to correct first phase area estimates for errors in classification and for actual changes in land use since the base photos were taken. Both procedures utilize standard photographic products to determine cover type acreages and boundaries, disturbance, land use changes, and plot locations. Subsequently, approximately 8500 of the photo plots were visited on the ground to establish plots for estimating cover type characteristics such as timber volume.

The second type of survey, (Phase II) is conducted by the Minnesota Department of Natural Resources (DNR) in cooperation with county and federal agencies, and industry. These surveys use aerial photography to map individual stand boundaries and describe stand age, density, size class and site quality for management decision-making regarding harvest, protection, regeneration and other concerns. An underlying interest in this project is to extend Phase II information to 100% coverage of the state.

The combined realities of shrinking public budgets and increasing information demands point to the need for automated image interpretation and easy access to computerized data bases which can be readily updated. Further, survey designs are needed which take advantage of digital image information. There is considerable and growing evidence which suggests that digital analysis of satellite-acquired data can be used to increase the efficiency and timeliness of forest inventory while reducing the overall cost. Further, the two types of inventory efforts may be combined at a considerable cost savings. This paper describes the rationale and design of such an approach.

Use of Landsat Data in Forestry

For many years foresters have effectively utilized aerial photography as a tool to help monitor and manage forest resources and aerial photographs are an integral part of most forest inventory procedures. However, the launch of Landsat-1 in 1972 added an entirely new dimension to the capability to acquire and analyze Earth resources data, and there has been much interest in the potential of satellite data and computer-aided analysis techniques to identify and map forest resources. In the Great Lakes Region, typified by relatively small stands of conifers, hardwoods, and mixed species, Landsat MSS classification accuracy at the species level has been unacceptably low (Mead and Meyer 1977, Roller and Visser 1980). Thus, MSS have not been widely used for forestry applications in this area.

It is important, however, to distinguish between results obtained using Landsat MSS data and the more recent results with Landsat Thematic Mapper (TM) data. A number of studies have shown that the information content of TM is considerably higher than that of MSS data and that the additional spectral bands and finer spatial and radiometric resolution of TM data result in significant improvements in classification accuracy for more specific information classes describing forest species (Horler and Ahern 1986) and forest stand characteristics (Peterson et al., 1986, Williams and Nelson 1986). In Minnesota, Moore and Bauer (1989) found a 15-20% increase in classification accuracy of TM data over MSS data, with TM accuracies of greater than 80% for seven and fewer species classes.

OBJECTIVES

The overall objective of the research is to develop and test procedures utilizing multispectral satellite data to inventory and map forest resources in the state of Minnesota. Specific objectives include:

1. Develop a methodology to use digital satellite data and computer-aided pattern recognition to classify forest cover types which will be compatible with and complementary to the other surveys conducted by the Minnesota Department of Natural Resources and the U.S. Forest Service.
2. Estimate forest areas and produce digital maps of the state's forest resources by species group at township, county and state levels, and determine the accuracy and precision of the forest area estimates and maps derived from satellite data compared to traditional forest inventory estimates.
3. Investigate alternative innovative approaches to satellite data classification, sampling and estimation.
4. Evaluate classification accuracy, area estimates, and type maps in relation to ecological and forest characteristics and in particular determine to what degree satellite data classifications can be used to obtain additional information such as stand density, size class, age, site quality, and disturbance.
5. Integrate the satellite data classification, sampling and estimation designs and procedures into the Forest Resource Assessment and Analysis Program of the Minnesota Department of Natural Resources.

The goal of the project is to develop and implement a methodology which will result in cover type area estimates of ±5% at the 95% confidence level at the state level and ±10%, 90% confidence at the county level. Other performance goals (criteria) of the final inventory procedure include: cost $.01-.02 per forested acre, one year to acquire and analyze the data, procedures can be implemented by the DNR with reasonable personnel and capital costs, and the system is flexible enough to meet changing conditions or requirements.

Two important underlying premises of the objectives and approach to be tested in the investigation are: (1) the synoptic view of Landsat provides the opportunity to obtain forest inventory information over large areas (i.e. state) and (2) by using computer data analysis methods to classify pixels distributed over counties and unique sampling designs, it is also possible to make accurate and precise estimates for local areas (i.e. counties). This approach offers a means to improve upon the sampling methods now used for making area estimates from ground-based surveys. At the same time the ground data (which are also used to estimate other parameters such as timber stand characteristics and therefore its collection cannot be abandoned) will be used to remove the bias from the satellite-based estimates.

TECHNICAL APPROACH

Multispectral satellite data, primarily Landsat TM, together with ground-based forest inventory data are being used to produce a current inventory of Minnesota forest lands, along with a methodology for frequent updates. The following sections describe the satellite data classification, sampling, and estimation procedures which will be the subject of research and development, implementation, and evaluation. The emphasis is on describing new approaches to forest inventory using satellite data while methods which are now standard procedures in the remote sensing field are described only briefly.

Sampling Design

In considering alternative designs we are primarily interested in collection of reference data for classifier training and adjustment for classification errors and bias. Large cluster designs have been shown to be an effective means of collecting forest inventory information (Scott et al., 1983). In particular, Bennesslah (1985) has shown that such layouts have decided advantages for ground checking of remotely sensed data. Advantages include ease of field work, variance reduction, and the provision of area data in terms of proportions rather than binary (0-1) counts. Proportion data facilitate the use of remote sensors because they can be relatively insensitive to scale problems. The design contemplated for current area and per unit area statistics by cover type is a single phase design involving remote sensing-based classification of the entire area and ground checking of a sample of large clusters. This would exploit the synoptic coverage of satellite-acquired digital remote sensing information by including all pixels in the population when addressing area estimation. A sample of clusters on the "imagery" would be observed on the ground to assess forest type. The clusters will be 10 to 100 acres in size. The ground sampling would establish an accurate type map for the cluster, i.e. boundaries and labels according to standards of interest to management. Clusters distributed systematically (with a random start) across the survey unit would then serve as the basis for training the digital image classifier.

The use of a large cluster allows training for both pixel classification and boundary identification. This is important for classification to achieve realism and associated type map accuracy. Importantly, the ground clusters are very inexpensively mapped. In fact, that might be done with large scale video or photographic imagery or by simple helicopter checking of digitally produced maps. The ground sample cluster result might also serve for iterative classifier adjustment to achieve ground estimated mean area estimates for maps over the entire survey unit.

Statistically one may view the cluster as a large and concentrated number of image classification plots. Their precision overall will be less than the same number of plots distributed randomly over the survey unit, but the cost of assessing them and particularly the realism they add to classifier training plus their low ground assessment cost should provide significant gains for the survey design. We see cluster mapping as costing little more than current ground plots (10 point variable radius plot clusters on an acre spaced every 1000-3000 acres). However, considering that a 20-acre cluster is the area equivalent of 20 current one-acre area classification plots, we believe considerable cost savings are possible.

Ultimately, current per unit area statistics by cover type will be drawn from small ground sample plots located within these clusters. Thus, clusters that are visited on the ground will contain traditional sample plots. These plots would be the basis for assessing growth, mortality, and to some extent removals and treatment opportunity. We suggest the ground observations follow design recommendations by Scott et al. (1983). They would also be the basis for forest growth models to project future supply.

Improving the ground information with the addition of data from previous surveys for the same forest types (not necessarily the same area, but the same forest types) is relatively straight forward. The weight given this information is yet to be decided, but it could be nearly equal to that of new plot information. Thus, where previous ground surveys are available one might have as many as two *n* rather than *n* ground plots to work with to estimate cover type or strata means.

The design of this survey involves a number of steps. Among these is developing the cost of clusters of various sizes as well as the correlation of remote sensing data with ground information. The development of procedures to integrate remote sensing classification with the information available from a type mapped cluster is also essential. With this basic cost, correlation and variance information, we are developing a simulator to study design alternatives. Then we will move to a large area trial for the aspen-birch survey unit of northeastern Minnesota. Fortunately, that unit was the subject of a reinventory during 1987-89. Thus, ground data from an existing survey are available to assess cost and classifier accuracy and "test" the design including a comparison with the existing FIA procedure.

The design or survey model will ultimately be refined by considering cost and variance submodels. The cost model will identify and detail cost components such as: design, training, administration/supervision, materials, travel, measurement, data entry, editing, compilation, estimation analysis, and overhead. The variance model will attempt to identify the sources of variation for each of the major variables in the survey and to develop parameter values for studying alternative designs.

Classification of Satellite Data

The training and classification procedures that will be tested in this project include an unsupervised clustering procedure followed by a maximum likelihood classification. Clustering will be done over areas which have been previously stratified into relatively homogeneous regions by examination of the satellite imagery in conjunction with soil and vegetation maps. Spectral cluster classes will be labeled using information from the ground clusters. In addition, we are researching the possibilities for combining a texture feature from 10m SPOT with the spectral features of Landsat TM data. An empirical Bayes approach (Morris 1986) is being considered as a means to incorporate spatial and prior probability information into the classification.

A necessary step in the classification of remotely sensed satellite data for forest inventory is the combination of statistical information from forest sampling units with that generated by the classification of the satellite data. How this can be done in a statistically sound manner is one of the research topics of this project. The general approach is that the statistical information derived from the ground cluster samples will be integrated into the classification process either as prior probabilities, or in a post classification procedure, or as some combination of the two. The forest sampling data will be used in two ways: 1) as an areal adjustment of the classification results for the different cover types in a predetermined stratified region, and 2) to provide co-occurrence information as to the relative likelihood that species are found adjacent to one another. The areal adjustment will be made on those pixels whose probability of classification is low or whose probability ratio with respect to the next most likely class is close to one. Probabilities will be generated by a maximum likelihood classifier that uses the spectral information from the Landsat TM data as input.

Area Estimation

Satellite data have commonly been used in agricultural surveys to increase the precision of ground-based estimates of crop areas. The bias involved is accepted as a concession to the extensive coverage provided by the satellite data. Normally, a regression or ratio estimator is applied with the notion being that the bias inherent in the satellite estimate will be removed by the estimator (Chhikara et al., 1986). A number of studies have shown that such an approach is satisfactory, though it does have problems (Heydorn and Takacs 1986). Perhaps a statistically cleaner way to view the combination of satellite and ground reference data is as a problem of calibration. In symbols:

$$Y_i = b_0 + b_1 x_i + \xi_i$$
$$X_i = x_i + \nu_i$$

where Y_i is the satellite data, X_i is the ground data, x_i is truth, the b's are parameters, and ξ and ν are errors. Combining these two equations gives the system which is actually observed:

$$Y_i = b_0 + b_1 X_i + (\xi_i - b_1\nu_i)$$

Estimating the unknown parameters of interest (the b's) by ordinary least squares (the usual approach taken) is thus seen to be biased and inefficient. Fuller (1987) provides a number of alternative estimation procedures that are more appropriate under the stated conditions including adjustment and grouping methods. Li (1988) has shown that for problems of a type related to area estimation using satellite and ground data, regression of X on Y may actually be preferred. The various calibration approaches will be evaluated in the context of the present study. In addition, variance estimation procedures will be developed and tested.

Of greater interest than broad scale estimation is the question of how best to obtain area estimates for geographic subdomains of the population. For example, in a state-wide inventory users will ultimately wish to have estimates of cover type acres for individual counties. Straightforward application of the regression or calibration procedures applied at the state level is possible but will give rise to large errors due to the paucity of data at the county level. Burk and Ek (1982) have shown through simulation study that careful grouping of subdomains allows application of shrinkage estimators directly to the cover type area estimates. However, such a procedure does not lend itself that well to complete automation of the estimation process. An alternative is to shrink the regression parameters themselves. Even though there may be disagreement whether two counties have similar cover type acreage, less argument will ensue over the supposition that the same regression adjustment (or calibration) can logically be applied in two counties. In symbols:

$$\mathbf{Y}_{ij} = \mathbf{X}_{ij}\,\beta_i + \xi_{ij}$$

where the subscript i refers to counties, the subscript j to samples within the counties, and the bold indicates vectors. The β_i's could then be assumed to have a common mean, and improved county parameter estimates could be had through the use of

$$\tilde{\beta}_i = \bar{\beta} + (\mathbf{I} - \mathbf{W})(\hat{\beta}_i - \bar{\beta})$$

where the bar indicates a common mean, the hat represents the usual regression or calibration estimator, and **W** is the James-Stein shrinkage coefficient (Burk and Ek 1982). An alternative approach to this same problem is presented by Battese et al. (1988) who employ an error components model and mixed model estimation procedures. In their approach the commonality between counties is accessed through error structure. These two approaches to subdomain area estimation and their associated estimation procedures will be compared.

Accuracy Assessment

A key component of the project will be to thoroughly test and evaluate the various area estimates and map products. The evaluation of results will be carried out at three levels as follows. The first step will be to assess the accuracy of the Landsat TM classifications. The discrete multivariate analysis techniques described by Congalton et al. (1983) will be used to evaluate classification error matrices. Secondly, the classification maps will be evaluated following procedures similar to those described by Rosenfield and Melley (1980). Finally the ground and satellite-based area estimates will be compared in terms of estimated bias, variance, and relative efficiency, and costs.

Project Evaluation and Implementation

Currently the DNR is conducting a remeasurement of the forest inventory permanent plots in Minnesota in cooperation with the USFS. It is also updating its stand assessment data. This public land data is being digitized as the base of the DNR's GIS system. These efforts, along with new statewide photo coverage, provide a firm base for evaluation of satellite data for forest inventory. The key questions for this assessment by the DNR are: 1. Can satellite data replace any of the traditional methods now used? 2. Are cover type, land use change and disturbance determinations from satellite data reliable enough for the statewide inventory program? 3. Can satellite data be used to update inventories of private non-industrial forest lands over which there is no control or existing reporting systems? 4. Can data interpretation and delivery functions be automated and linked to the DNR's GIS for better client service and cost savings? 5. Will the use of satellite data shorten the time period between inventory updates and reduce their overall cost while maintaining a high enough degree of reliability to answer the complex questions being asked about Minnesota forest resource availability and condition?

From the DNR perspective, the end result of this investigation should be an automated procedure to use satellite data to estimate county and state forest cover type acreage totals, summarize cover type changes (land use, harvest, catastrophe, disturbance) and capability to update statewide cover type maps in existing Minnesota forest inventory data bases for use by the DNR.

ACKNOWLEDGEMENTS

Support for the project has been provided by NASA Grant NAGW-1431 and the University of Minnesota Agricultural Experiment Station. The cooperation of the Minnesota Department of Natural Resources, Forestry Division, Resource Assessment Unit, is gratefully acknowledged.

REFERENCES

Battese, G.E., R.M. Harter and W.A. Fuller. 1988. An error-components model for prediction of county crop areas using survey and satellite data. J. of American Stat. Assoc. 83:28-36.

Bennesslah, D. 1985. Forest area estimation using cluster sampling in single and double phase sampling designs. Ph.D. Dissertation, University of Minnesota, St. Paul (University Microfilm #DA85-26459). 171 p.

Burk, T.E. and A.R. Ek. 1982. Application of empirical Bayes/James-Stein procedures to simultaneous estimation problems in forest inventory. Forest Science 28:753-771.

Chhikara, R. S., J. C. Lundgren, and A. G. Houston. 1986. Crop acreage estimation using a Landsat-based estimator as an auxiliary variable. IEEE Geoscience and Remote Sensing 24:157-168.

Congalton, R., R. Oderwald, and R. Mead. 1983. Assessing Landsat classification accuracy using discreet multivariate statistical techniques. Photogrammetric Engineering and Remote Sensing 49:1671-1678.

Ek, A. R. 1983. Use of forest growth models in management inventories. Proc. IUFRO Subject Group S4.02, meeting in Finland, Sept. 5-9. Published by Univ. Helsinki, Dept. Forest Mensuration and Management, Research Notes No. 17, pp. 129-136.

Fuller, W.A. 1987. Measurement error models. John Wiley and Sons. New York. 440 pp.

Hansen, M. H. and T. E. Burk. 1986. Testing the consistency of forest cover type estimates across survey periods. In: Use of auxiliary information in natural resource inventories. Society of American Foresters Publications No. 86-01.

Heydorn, R.P. and H.C. Takacs. 1986. On the design of classifiers for crop inventories. IEEE Geoscience and Remote Sensing 24:150-156.

Horler, D.N.H. and F. J. Ahern. 1986. Forestry information content of thematic mapper data. Intl. J. of Remote Sensing 7:405-428.

Jakes, P. J. 1980. The fourth Minnesota forest inventory: area. USDA Forest Service, Resource Bull. NC-54, North Central Forest Experiment Station, St. Paul, Minnesota.

Lewis, B. J. 1984. An overview of Minnesota's forestry sector. Proc. Governor's Conference on Minnesota Forest Industry as a Force in Economic Development. Minneapolis, Minnesota.

Li, Chuanzhong. 1988. Mathematical models in forest resource management planning. Swedish University of Agricultural Sciences. Report number 18. 197 pp.

Mead, R. and M. Meyer. 1977. Landsat digital data application to forest vegetation and land use classification in Minnesota. Proc. Machine Processing of Remotely Sensed Data Symposium. Purdue University, W. Lafayette, Indiana. pp. 270-279.

Moore, M.M. and M.E. Bauer. 1989. Classification of Forest Vegetation in North-Central Minnesota Using Landsat Multispectral Scanner and Thematic Mapper Data. Forest Science, in press.

Morris, C.N. 1986. Binary image classification. In: Computer Science and Statistics: Proc. 18th Symposium on the Interface. Amer. Statistical Assoc., Washington, D.C. pp. 71-74.

Peterson, D. L., W. E. Westman, N. L. Stephenson, V. G. Ambrosia, J. A. Brass, and M. A. Spanner. 1986. Analysis of forest structure using thematic mapper data. IEEE Geoscience and Remote Sensing 24:113-121.

Roller, N. E. and L. Visser. 1980. Accuracy of Landsat forest cover type mapping in the Lake States region of the U.S. Proc. 14th Intl. Symposium on Remote Sensing of Environment, San Jose, Costa Rica. pp. 1511-1520.

Rosenfield, G. H. and M. L. Melley. 1980. Applications of statistics to thematic mapping. Photogrammetric Engineering and Remote Sensing 46:1287-1294.

Scott, C.T., A.R. Ek and T.R. Zeisler. 1983. Optimal spacing of plots comprising clusters in extensive forest inventories. In: Renewable Resource Inventories for Monitoring Changes and Trends. College of Forestry, Oregon State University, Corvallis. pp. 707-710.

Williams, D.L. and R.F. Nelson. 1986. Use of remotely sensed data for assessing forest stand conditions in the Eastern United States. IEEE Geo. and Remote Sensing 24:130-138.

Inventory of the Jarrah Forest of Western Australia Using Large-Scale Photography and Ground Sampling.

Paul Biggs
Manager, Ground & Air Surveys
Department of Conservation and Land Management
P.O. Box 733, Bunbury, 6230, Australia

ABSTRACT

Large-scale aerial photography and ground samples were combined in a two-stage inventory of 1.25 million hectares of jarrah (*Eucalyptus marginata*) forest in Western Australia. The large-scale photography (1:1000) scale was acquired using twin 70 mm cameras mounted on a 7.5 m transverse boom on a Bell Jetranger helicopter. A satellite-based Global Positioning System (GPS) was used to navigate along flight-lines and was linked to a lap-top computer which fired the cameras on a fixed distance interval and recorded the position of every photograph to ± 25 m. Photo pairs were measured using a small format analytical stereo digitizer. Ten percent of all photo samples were measured on the ground to derive estimates of volumes of log products. The position, size and severity of every defect in every sample tree was recorded in detail, then a computer program was used to predict the utilization of the tree under a given set of utilization standards. New estimates could be derived for different utilization standards by reprocessing. GIS window generation and point-in-polygon techniques were used to assist estimation of volumes for any unit of forest. Data from photo samples and ground samples within a window were extracted from an INFO database and passed to a data processing program for generation of volume estimates. The major feature of this inventory was its flexibility, allowing volume estimates to be generated quickly for different areas of land and under different utilization standards.

INTRODUCTION

The jarrah forest covers two million hectares in the south west of Western Australia. Jarrah (*Eucalyptus marginata* Donn. ex Sm.) provides a valuable timber for structural and decorative uses while marri (*E. calophylla* R.Br. ex Lindl.), the other main species in the forest, is suitable for high quality paper manufacture. An area of 1.3 million hectares is available for timber production.

The inventory described in this paper has been designed to provide detailed information about the jarrah forest resource in the shortest possible time, while maintaining the flexibility needed to examine different management options. The design has been constrained by several factors.

Intensity

The sample needs to be intense enough to provide estimates of volume for many different products over any portion of the forest greater than 10 000 ha. This is to facilitate planning for any one of the several sawmills or other forest industries located in the region.

Timing

The inventory is to be completed in only three years to allow a review of cutting levels in 1991.

Flexibility

The jarrah forest now is a mixed-age regrowth forest, resulting from over one hundred years of logging. As the traditional oldgrowth logs become scarcer, existing sawmills must necessarily convert to cutting the smaller regrowth logs. Research and investment in new technology will result in a gradual change in the type of logs accepted by the mills in the future. Estimates from the inventory of volumes of products cannot to restricted to current assessment of log grades but must be able to cope with change.

Areas

Volume estimates from the inventory cannot be linked to any particular land base. For example, new industries are being encouraged to make the silvicultural treatment of regrowth jarrah more economic. The supply zones of these industries will not be finalised until the economics of each industry are more certain. The land base may also alter with the reservation of new areas for purposes other than timber production or with the progress of harvesting. This paper contains a description of the some of the techniques which were implemented to allow the inventory to satisfy these constraints.

GENERAL DESIGN OF THE INVENTORY

The inventory used two-stage sampling, with large-scale aerial photography in the first stage, to estimate gross bole volume quickly and cheaply, and ground measurements in the second stage to estimate volumes of log grades and to correct for bias in the photo estimates. The photography was acquired at a scale of 1:1000 using twin cameras mounted on a transverse boom on a Bell 206B helicopter (the method described by Lyons (1966)). The aerial photography provided a relatively intense sample (0.24%) over large areas in a short time. The height of each tree taller than 10 m and within 20 m of the centre of overlap was measured. Tree volumes were estimated using a height - volume function derived for the species involved. The function for each species was of the form

$$\ln(\text{volume}) = a + b \ln(\text{total height}).$$

Ground measurements were made on ten percent of the photo sample plots. The ground measurements were used to estimate the timber grades and to correct for any bias in the photo estimates of bole volume. Samples from both stages were selected systematically. The photographs were acquired at 500 m intervals along flight lines 1000 m apart. Ground samples were selected as every fifth sample plot on every second flight-line, giving a spacing of 2000 m by 2500 m.

INNOVATIVE COMPONENTS

GPS Navigation

Global Postioning System (GPS) navigation was used when acquiring the aerial photography. GPS navigation instruments generally are capable of determining the position of a vehicle to ± 25 m every second, using signals from the constellation of "NAVSTAR" satellites.

For this application, a Trimble 10X GPS navigator was connected, via an interface unit, to a laptop computer and the two cameras. The Trimble 10X directed the pilot along pre-programmed flight lines.

Simultaneously, it provided position data every second to the laptop computer. The computer calculated the distance flown since the last photograph and then fired the cameras when that distance reached 500 m ± 15 m. As each photograph was taken, the computer recorded the position, time, satellite status, frame number and the altitude (from a radar altimeter).

Back on the ground, the location data were transferred from the laptop computer to the Department's Geographic Information System. Using this, maps were produced at 1:25,000 showing the location of every photo sample. These maps were used by the field crews to locate the sample points which required ground measurement. The precision of the maps was affected by the precision of the GPS, the base map and the overlay process. The resulting precision was ± 75 m, which was sufficient for the field crews to find the exact plot centre by viewing the photos on a portable light board

The GPS navigation system had many benefits. It enabled the accurate placement of flight lines and uniform spacing of photopoints, irrespective of the air speed of the aircraft. This ensured that the sample distribution was uniform.

It also provided digital data on the locations of each photo-pair, allowing quick production of location maps by computer aided mapping.

Modified Hasselblad 500 EL/M camera.

Adam MPS-2 small format analytical stereo digitizers were used to measure the tree dimensions on the aerial photography. Due to the distortion caused by their focal plane shutters, the photography acquired using the Department's Vinten reconnaissance cameras could not be measured using these analytical instruments.

Three cameras with "between-the-lens" shutters were investigated as alternatives. The Rolliflex 6006 Metric was found to be unsuitable because large capacity magazines were not available. The Hasselblad MK70 was considered too expensive.

The third choice, the Hasselblad 500 EL/M had the same type of shutter as the MK 70 and large capacity magazines were available (second-hand). However these cameras had no registration marks (fiducials) to use when completing an iinterior orientation of the photographs in the analytical plotter. Neither could the edges of the format be used because these were actually edges in the magazine, which could change position with respect to the lens.

The solution was to install reseau plates in the bodies of two Hasselblad 500 EL/M cameras. The reseau plates were glued into a frame fabricated out of aluminium so that the marked surface of the glass plate almost touched the film. The resulting marks on the film were clear and sharp and worked well as fiducial marks in the plotter. The plate also improved the film flatness during photography.

There were two drawbacks with the modified cameras. The introduction of the glass plate altered the focussing of the lens, although this had no serious effect unless photographing very close objects (30 m) with a narrow depth of field. Secondly, the plate had to be removed when a camera jammed in mid-cycle. If this occured, then a new lens calibration was needed.

Assessment of wood qualities.

A new assessment technique was developed, independantly of the aerial photography, for estimating the grades of timber in each sample tree. This was critical because only ten to twenty percent of timber in the jarrah forest may be suitable for sawlogs.

The technique was designed to allow estimates of volumes of wood grades to be made at any time under any set of utilization standards. Instead of the field crews assessing the grades in each tree under the set of standards for the day, the crews assessed the type, severity, extent and position of every defect or "quality" which could affect the utilization of the wood in the tree.

A computer program was then used to sort through the data and determine which grades could be cut from each tree. To do this, the relevant set of utilization standards (minimum and maximum log dimensions and acceptable defects) were entered to the program as well as a priority order for the sorting procedure. The program could be re-run to estimate volumes of log grades available under different standards, or even volumes of completely new classes of logs.

This method should prevent the inventory data from becoming outdated and reduce the amount of measurements required in the future.

GIS assisted processing.

To achieve the required flexibility in the areas for which volume estimates could be given, the photo plot data and ground plot data were stored in an "INFO" database and linked to the location data from the GPS navigation system.

At any time, an area could be defined for which an inventory statement was required, using the Department's Geographic Information Systems. Any sample plots which fell into this area could be identified using the ESRI "point-in- polygon" software and the data from these plots retrieved from the database and sent to statistical software for processing. The plot selection could be re-run and a new volume statement derived in the event that the area base changed, for example due to an increase in the supply area, cutting of some of the forest or the declaration of new reserves.

CONCLUSIONS

New techniques have been implemented to enable the inventory of Western Australia's jarrah forest to meet the objectives of speed and flexibility. GPS navigation has enabled sample aerial photography to be obtained at a uniform intensity over all areas. It has also allowed any of the photo samples to be relocated efficiently for ground measurement. The installation of reseau plates in Hasselblad 500 EL/M cameras has provided semi-metric cameras with large capacity magazines much cheaper than the alternatives, although difficulties persist with their operation.

A new method of assessing wood qualities instead of log grades in trees has allowed the estimation of volumes of log grades to be made at any time under any set of utilization standards.

Finally, the use of a Geographic Information System to help store and retrieve sample data has provided the possibility of generating estimates of volumes for any arbitrary area of land, independant of any pre-conceived sampling framework.

ACKNOWLEDGEMENT

Aspects of this project form part of the author's Ph.D. research at the University of Melbourne, Australia.

REFERENCES

Lyons, E.H., 1966: Fixed air-base 70 mm photography, a new tool for forest sampling. Forestry Chronicle, vol. 42, no. 4, pp 420-431.

A Practical Example of Forest Mapping and Timber Volume Estimation Using Landsat-5 TM Imagery*

David Chung Liang Lee, René Antonio Novaes
Pedro Hernandez Filho, Flávio Jorge Ponzoni
Researchers of Instituto de Pesquisas Espaciais-INPE
Secretaria Especial da Ciência e Tecnologia
Caixa Postal 515, 12201, São José dos Campos, SP, Brazil

ABSTRACT

The objective of this project was to obtain basic forest information for Uruguay. This report describes a methodology which used Landsat TM data as a first segment of a two-stage sampling plan designed to produce typical forest inventory information for plantations (pine and eucalypt) and to map forest areas. From all the results obtained, the project could conclude that Landsat-5 TM imagery, at a scale of 1:100,000, served well as the primary source of data for forest area mapping and timber inventory over a nation-wide level.

INTRODUCTION

The main objective of this project was to obtain basic forest information for development, promotion and utilization of alternative sources of energy based on the use of the country's forest resources, and to evaluate the forest Landsat Thematic Mapper (TM) data as an aid to forest plantations management in Uruguay. The project attended to the need of evaluating Uruguay's forest potential in order to develop a reforestation programme for the use of wood as an alternative source of energy. To accomplish this evaluation, it was necessary first to establish the spatial distribution of the forest cover and evaluate the existing volume of stand timber. Considering that the Landsat TM imagery covers the same area every 16 days and has spatial resolution of 30 meters, these data, added to the ground information surveys, provide a cost-effective means of evaluating current timber resources and planting areas, at a national level. The specific objectives of the project were: (1) to obtain maps of Uruguay forested areas at scales of 1:100,000, including natural stand with 50% canopy; (2) to obtain information on forest plantations for those maps; (3) to generate a map of the forested areas of Uruguay at a scale of 1:1,000,000; (4) to compute the stand timber volume in plantations areas. The area under study in this project corresponds to the Uruguayan territory, which is composed of 19 states called "Departamentos" and represents 176,027 square kilometers. Since the available recent aerial photographs do not cover all the country, it was necessary to divide the territory in two zones: Zone I and Zone II. Recent aerial photographs were available for Zone II only.

METHODS AND MATERIALS

Visual Image Interpretation

The territory of Uruguay is covered by 44 quadrants of Landsat-5 imagery. The material used for interpretation and map compilation were black and white paper print images from Landsat TM, band 3 (0.63 to 0.69 μm), band 4 (0.76 to 0.90 μm) and some band 5 (1.55 to 1.75 μm) at 1:100,000 scale and band 3 at 1:1,000,000 scale. The methodology used was based on the visual interpretation of unenhanced Landsat TM imagery. The image analysis was conducted in two parts. The first part was conducted previously to the ground checking, and the second analysis was conducted after the field work, when the interpreters were more familiar with both ground features and their correspondence in the imagery. Based on spectral and spatial image attributes the interpretation was made by using a visual methodology. Taking into account the image spatial attribute it was analyzed the pattern and size of the topographical site occupied by the artificial (pine and eucalypt) and natural forest. Relating to the image spectral attribute it was analyzed the texture and tonality.

---

* This is a condensed version paper of report for Project No. DP/URU/83/013, Contract No. 85/125/MK; "Assessment of Forest Resources in Uruguay" submitted to the United Nations Industrial Development Organization (UNIDO).

Thematic Map Compilation and Area Measurement

The interpretation of a Landsat image was drawn with a 5 km grid at 1:100,000 scale. The base map sheet was also used to control the organization and the localization of the information extracted from the images. The scale variation of the images and errors on the maps produced discrepancies between forest boundaries and map features. To minimize this kind of error, in the order of one kilometer on the ground at most, a procedure was adopted to distribute it over a whole scene. It consisted of working one by one each adjacent cell (5 km grid) of the base map and performing small adjustments between the interpretation overlay and the base map. To measure the forest areas it was used the overlay containing the interpreted forest classes. The area of plantations was estimated by superimposing a 0.4 mm grid overlay, and for estimating natural forest a 1 mm grid was used. Area acreages by class of Primary Sampling Units (PSU's) were then measured and listed. For Zone II, a Secondary Sampling Unit (SSU) as being 1/625 of one PSU was defined. The area counting inside each SSU was done with a 1 mm grid overlay but superimposed on a mosaic of aerial photos at a scale of 1:20,000. Each photomosaic has 25 x 25 cm to cover a PSU, and each 1 x 1 cm corresponds to one SSU (200 x 200 m area on the ground).

Sampling Design

The methodology and procedure for timber volume estimate in multistage inventory with variable probability has been used in Brazil (Lee, et al., 1983, 1984 and Hernandez Filho, et al., 1985). With this experience, one established a two-stage sampling design with probability proportional to size (PPS) and simple random (equal probability) sampling. These procedure were used to estimate the plantation stand timber volume in the two zones.

Sampling Scheme. For Zone I, the method consisted of a two-stage sampling design using PSU's selected with PPS. In the second stage, each SSU was chosen on a simple random basis. The unbiased estimator (Cochran, 1977, P. 306-308) of the total timber volume (V_t) of each class can be expressed as:

$$V_t = \frac{1}{m} \sum_{i=1}^{m} \frac{N_i}{P_i n_i} \sum_{j=1}^{n_i} V_{ij} ; \qquad P_i = x_i / \sum_{i=1}^{M} x_i$$

where: P_i is the conditional probability of drawing the ith PSU, V_{ij} is the field measured value of timber volume in the jth SSU of the ith PSU, x_i is the forest area of a class in the ith PSU, N_i is the number of SSU's in the ith PSU, m is the number of PSU's selected in the class, M is the number of PSU's in the zone per class, n_i is the number of SSU selected in the ith PSU. The estimator for the variance is:

$$\text{Var } V_t = \frac{1}{m(m-1)} \sum_{i=1}^{m} \left(\frac{V_i}{P_i} - V_t\right)^2.$$

For Zone II, the variable probability sampling design was used where sample units of two stages were all selected based on PPS. The size was the acreage of plantations estimated within the PSU and SSU units. A generalized model for a two stage estimator based on Langley (1975) is:

$$V_t = \frac{1}{m}\sum_{i=1}^{m} \frac{1}{P_i n_i} \sum_{j=1}^{n_i} \frac{V_{ij}}{P_{ij}}; \qquad P_{ij} = x_{ij} \Big/ \sum_{j=1}^{N_i} x_{ij},$$

where: P_{ij} is the conditional probability of drawing the jth SSU in the ith PSU, x_{ij} is the forest acreage of the jth SSU in the ith PSU. The unbiased estimate of sample variance can be obtained from the estimate of the first stage, that is:

$$\text{Var } V_t = \frac{1}{m(m-1)} \left(\sum_{i=1}^{m} \frac{V_i^2}{P_i^2} - mV_t^2 \right).$$

This estimator of variance, based on a simple stage estimate, may also be used to determine the unbiased estimate of sample variance for a two-stage sample estimate (Langley, 1975). The precision of variable probability sampling is dependent on the relationship between the predicted values, which determine the sampling probability, and those determined by precise measurement (in this project it was the acreage of forest land areas). The premise here is that the greater the forest area within prescribed sample units, the greater the timber volume.

Sample Size and Allocation. The sample size is calculated by: $m = t^2 \cdot CV^2/E^2$, where: m is the sample size in the first stage, t is the normal deviation (two-sided) for a given confidence level, CV is the coefficient of variation of estimator, E is the allowable error in percent. The basal area per hectare of forest land (per class), estimated from a selected sample in the presampling field work, provides an approximation of the variation in forest land over the area being sampled and can be used to estimate the size of the sample in the first stage. As to sample allocation, the selection probabilities used in this project were equal probability and PPS. This procedure allows giving to the large units higher selection probabilities, since they represent larger portions of the population. This was the rationale for PPS selection which adapted for the project. Both methods were performed with replacement.

Field Volumetric Measurement

In the second stage, the field work was carried out to determine the stand timber volume for the selected sample units (SSU's). The mean stand volume, mean basal area and mean stand height from the sampled plots were estimated using the Strand's Line Sampling method with the Relaskop of Bitterlich (Bitterlich, 1984, P. 43-46). Using the field work data, the SSU's, the PSU's and the total stand timber volumes were then calculated using the equations defined in before section.

Accuracy Assessment

The errors in the interpretation of the Landsat imagery affect: (1) the placement of the boundaries, and (2) the identification of the classes delimited on them. This leads to errors of misclassification and misplacement and, therefore, to errors of area estimation for the classes concerned. To calculate the accuracy of the forest cover mapping, a sample from the total population of the classified units (SSU's) was selected and a field check for this chosen sample was carried out.

RESULTS AND DISCUSSION

Visual Interpretation

Table 1 shows a photointerpretation key for the three main classes using Landsat TM imagery. But tone and texture of these classes vary somewhat according to the age, site, density, crown size and type of timber growth. This experience has permitted to observe that:

1. On image prints, the pine plantation appears in dark grey (low reflectance) on band 3 (chlorophyll absorption band), but it appears lighter (high reflectance) on band 4 (reflecting IR radiation band) than on band 3.

2. Shape and size were a very important key to distinguish plantations and natural forest. In flat pastureland, the isolated plantation areas areas can easily be determined at the minimum area of one hectare. The shape of plantation areas always shows regular forms.

3. Shadow is useful for identifying plantations, specially for isolated small areas in band 4. These areas are always accompanied by dark grey shadows.

4. Fairly homogeneous plantation areas of more then 4 hectares can be identified on image of band 3, in part due to their linear boundaries. Separation of pine and eucalypt is more clear on band 4.

Table 1 - Photointerpretation key of analysis

	Tone		Texture	
Class	Band 3	Band 4	Band 3	Band 4
Pine	DG (dark grey)	G (grey to DG)	F (fine)	MC (more coarse)
Eucalypt	G to DG	Light Grey	F	F
Nat. Forest	G to DG	G	MC	MC

Interpretation and Mapping Accuracy

The confusion matrix was generated through visual interpretation results. They display the corresponding accuracies which were evaluated when compared with ground data from a field sample check. Table 2 shows the overall interpretation and mapping accuracies which were 90.6% and 83.3% respectively. Thus, the interpretation accuracies for pine and eucalypt are 82.4% and 95.5% respectively, and the mapping accuracies for pine and eucalypt are 81.9% and 81.4% respectively. The area of eucalypt plantations had been overestimated by 15.3% and that of pine plantations had been underestimated by 17.6%. These are somewhat low accuracies in comparison with the commission error of pine (0.7%) and the omission error of eucalypt (4.5%) plantations. The confidence intervals and the sampling errors of interpretation and thematic mapping accuracies were calculated and each of them at 95% confidence level. All accuracy requirements for each class of interpretation and thematic mapping were met (sampling error around 5%) except those for natural forest (Table 3).

Forest Area

Uruguay has a total area of 176,027 square kilometers, and according to this project we conclude that the total forest area is about 3% of the total land area in Uruguay. The plantation forest (pine and eucalypt) is about 23.7% of the total forested area. Natural forests are concentrated basically along rivers and drainage. In the pastoral region, part of plantation forest types were not planted for the purpose of timber woods, but for sheltering livestock in stockfarms or to serve as windbreak belts. This resulted in small and sparse forest plantations with sizes from 0.25 to 20 hectares.

Volumetric Results

The statistical results of the total stand timer volume for Zone I and II are presented in Table 4. The form factor used for pine plantations was 0.45 and for eucalypt plantations was 0.50. The stand timber volume per "departamento" in cubic meter was generated by the data of forest acreage and the mean stand timber volume. The total volume timber of Pinus spp. and Eucalyptus spp. plantations in the whole country were estimated as 3,539,418.15 cubic meters and 24,231,118.83 cubic meters respectively.

Table 2 - Confusion matrix for accuracy analysys

	Field check			
Image interpretation	Pine	Eucalypt	N. Forest	Others
Pine	267	2	0	0
Eucalypt	49	276	0	1
Natural Forest	0	0	69	6
Others	8	11	7	197
Total	324	289	76	204
Comission error (%)	.7	15.3	8.0	11.7
Omission error (%)	17.6	4.5	9.2	3.4
Interpretation (%)	82.4	95.5	90.8	96.6
Mapping accuracy (%)	81.9	81.4	84.2	85.7
Overall interpretation accuracy (%)				90.6
Overall thematic mapping accuracy (%)				83.3

Table 3 - Accuracy statement

	Interpretation		Thematic Mapping	
Class	Sampling Error (%)	% Correct & Conf. Interval	Sampling Error (%)	% Correct & Conf. Interval
Pine	5.04	82.4 ± 4.2	5.13	81.9 ± 4.2
Eucalypt	2.50	95.5 ± 2.4	5.52	81.4 ± 4.5
Nat. Forest	7.20	90.8 ± 6.5	9.80	84.2 ± 8.3

Table 4 - The results of the estimated timber volume

Class	Pine	Eucalypt
Zone 1		
Total timber volume (m_3)	1,710,608.20	15,429,409.34
Standard error (m_3)	140,002.10	1,933,835.94
Relative standard error (%)	8.18	12.53
Upper 95% confidence limit (m_3)	1,985,012.32	19,219,727.78
Lower 95% confidence limit (m_3)	1,436,203.88	11,639,090.90
Relative confidence interval (%)	± 16.03	± 24.56
Mean & conf. interval (m_3/ha)	145.34 ± 23.30	243.64 ± 59.84
Zone 2		
Total timber volume (m_3)	1,828,805.95	8,801,709.49
Standard error (m_3)	199,320.85	855,837.62
Relative standard error (%)	10.90	9.72
Upper 95% confidence limit (m_3)	2,176,275.98	10,537,951.18
Lower 95% confidence limit (m_3)	1,438,140.98	7,065,467.71
Relative confidence interval (%)	± 21.36	± 19.06
Mean & conf. interval (m_3/ha)	225.38 ± 48.14	217.57 ± 41.47

CONCLUSIONS

Several conclusions were reached during this project and may be summarized as follows:

1. It was possible to establish that the forest area mapping and inventory at a national level in Uruguay can be done satisfactorily using Landsat-5 TM images.

2. The characteristics of forest classes may be recognized using combinations of tone, texture, shape and size in two spectral bands (3 and 4) of Landsat TM imagery.

3. Shadowing is an useful tool in the identification of forest plantations, specially for isolated small areas in TM band 4. These areas are always accompanied by dark grey tones.

4. Visual interpretation is a very simple and efficient methodology for mapping forest land using Landsat TM imagery at a scale of 1:100,000 in paper prints. This is indicated by the figures obtained for the overall interpretation and mapping accuracies of 90.6% and 83.3% respectively.

5. At 1:100,000 scale, the Landsat TM imagery shows differences in forest stand structure, usually in a conspicuous manner. Nevertheless species identification may sometimes be difficult, such as differences among species from Eucalyptus, Populus and Salix spp.

6. Generally the greater the forest area within a sample unit, the greater the timber volume. However, age cutting and the difference in growth rates lead to variations in volume not necessarily associated with the area size.

7. The Strand's Line Sampling method was useful only when a sufficiently large number of a sample size was used. When the trees not uniformly distributed in a SSU, the method employed could not be a good estimator of basal area, height and volume of a single SSU just a few measurement.

8. In future works it is suggested an attempt to establish a regression model between data obtained from satellite, aerial photographs and field work, for different forest classes, in order to analyze and improve the final results.

9. Accordingly to the objectives of this inventory, the results and the accuracy of the obtained information can be considered satisfactory. Nevertheless, for a more comprehensive analysis of the forest resources, it is suggested a plantation stratification by groups of species, sites and ages.

REFERENCES

Bitterlich, W., 1984: The Relaskop Idea. Commonwealth Agricultural Bureaux, Farnham Royal, England, PP. 43-46.

Cochran, W.G., 1977: Sampling Techniques, third edition. John Wiley and Sons, New York, U.S.A., PP. 306-308.

Hernandez Filho, P., Lee, D.C.L. and Shimabukuro, Y.E., 1985: Relatório das atividades do Projeto IBDF/INPE (Subprojeto reflorestamento) de 1982-1984. São José dos Campos, INPE, Brazil, (INPE-3396-RTR/067).

Langley, P.G., 1975: Multistage variable probability sampling: theory and use in estimating timber resources from space and aircraft photography. Ph.D. Dissertation, University of California, Berkeley, U.S.A.

Lee, D.C.L., Hernandez Filho, P. and Shimabukuro, Y.E., 1983: Inventário da floresta plantada através de amostragem em múltiplo estágio com probabilidade proporcional á grandeza. São José dos Campos, INPE, Brazil, (INPE-2869-RPE/441).

Lee, D.C.L., Hernandez Filho, P. and Shimabukuro, Y.E., 1984: Forest inventory using multistage sampling with probability proportional to size. In: International Archives of Photogrammetry and Remote Sensing. Proceedings, Rio de Janeiro, Brazil. Vol. XXV, part A7, Commission VII, PP. 285-292.

An Aerial Photographic Key as an Aid to Improving Photointerpretation Accuracy

Rachel Riemann Hershey
Research Officer, SERRL/LSE
Geography Dept., London School of Economics
Houghton St., London WC2A 2AE, UK

ABSTRACT

The application of large scale aerial photography for timber inventories is a combination of photointerpretation and photogrammetry; the former to determine species conditions and location, and the latter to determine the dimensions of trees, stands and areas. An aerial photographic key as an aid to photointerpretation can significantly improve interpretation accuracy.

A modular aerial photographic key to 21 New England forest cover types was developed on the basis of direct and indirect identification clues. Initial analysis determined the most appropriate emulsion, scale and season of photography for the identification of forest cover types, and suggested those direct and indirect clues interpreters used most often in the identification of tree species and cover types. The key illustrates each type with 1:6000 scale color infrared and 1:20000 panchromatic vertical stereograms, and provides text and graphical information in a common format about type composition and range, site preference, associated species, and identifying features. A test of the key using five interpreters and 95 test stands was employed to assess the accuracy with which the 21 cover types could be identified. The test showed that the key significantly improved interpretation accuracy. Assembly of the key is briefly discussed, covering the forest cover type system, the formats chosen, the components of the key, and the key arrangement.

INTRODUCTION

The translation of information from its pictorial format on the aerial photos into written form is called airphoto interpretation, and is considered correct if it agrees with what would have been recorded by an observer on the ground. The use of aerial photography is an advantage only if the interpretation is accurate enough to reduce substantially the amount of data gathering necessary in the field. In the field of forestry, where the areas of interest are often not easily accessible from the ground, aerial photography has been widely used. Conducting forest surveys, mapping, investigating damage, arranging timber sales, and locating roads, trails, telephone lines and firebreaks can all be accomplished with the help of aerial photography. Accurate aerial interpretation, however, especially of tree species and forest cover types, requires experience. There is enough information available on an aerial photo for such identification, but experience is required to accurately assess it all. A key to airphoto interpretation can contain information and example stereo photography to describe distinguishing features, demonstrate correct identification and generally increase an interpreter's familiarity with the subject. A good key can be a training tool for the new interpreter, a guide for the amateur, and a refresher and reference for the experienced interpreter.

METHODS

Formats Chosen

Classification system. The classification system used in a key must reflect that which is used in practice. The Society of American Foresters' Forest Cover Types of the US and Canada (Eyre 1980) was chosen as the basis for the classification system used in this key. It is an example of an association type, defining 90 eastern and 55 western forest cover types in terms of composition, geographic distribution, ecological relationships including site preference, and associated vegetation. Although this system represents neither a fully articulated scheme of phytosociology nor an ordered hierarchy of habitat types, the SAF type terminology is current and well-recognized among eastern foresters both at the scientific and the working levels. In addition, the types fit neatly into the US Forest Service Renewable Resource Evaluation

(RRE) type groups, on the basis of which decennial Forest Survey data are compiled nationwide (Eyre 1980). Because it offers these advantages, and in the absence of any generally accepted and widely recognized alternative, the SAF cover type system was chosen as the organizing principle for the projected key. A preliminary list of 27 types found in New England was drawn from the SAF guide from which 20 were chosen to be represented in the key.

Selection vs. Elimination Key. The guide to photo identification that was assembled is a selection key. The large number of tree species in New England, some of them displaying quite similar features in an image, makes an elimination key hard to devise for this area and rather limiting. The details chosen by a dichotomous elimination key to be the distinguishing features under one set of circumstances may not be visible if the photography is a different emulsion, season, scale or format. The added complexity of forest cover types, each of which has a range of compositions and thus a range of possible appearances, makes identification from single distinguishing details more difficult. Thus, for the identification of forest cover types, a selection key, which offers more complete illustrations and descriptions, is most appropriate.

Modular format. To allow for expansion of the guide to address additional cover types, other seasons, scales and types of photography, a modular, loose-leaf format was chosen. Further modification, limiting each type to the two sides of a single loose leaf page, was chosen to facilitate these incremental improvements and corrections, and allow single types to be removed or reorganized by users to fit the current project.

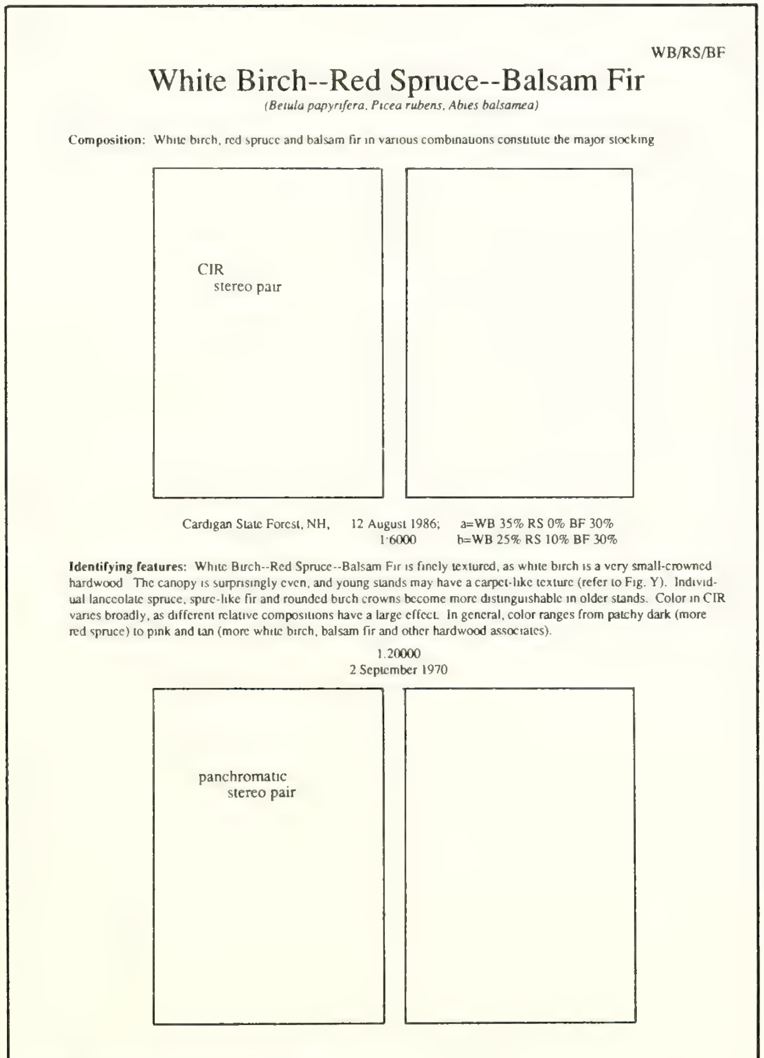
Photography and equipment. The 35mm format was chosen for the project because both camera and film were relatively inexpensive, and the equipment could be used without a mount and was much simpler to handle, especially from a small plane. CIR was chosen as the primary emulsion as it reveals a sharp distinction between hardwoods and conifers, and with the greater reflectance by vegetation in the near IR, adds substantial information to discriminate between hardwoods as well (Knipling 1970). The following list provides a summary of the equipment used.

camera--Nikon F2A with MD2 motor drive
lens--135mm Nikkor, f/3.5
filter--Wratten 12 (yellow or minus-blue)
film--Kodak Ektachrome infrared film 2236
airplane--Cessna 172 (hired)
supporting photography--ASCS 9" panchromatic (1:20000)

Components of the Key

Summary. Each key page includes the type name and definition, stereo photo examples of the type (1:6000 color infrared and 1:20000 panchromatic of the same area); a paragraph of identification features; comparisons with easily confused types; graphical representations of species site preferences, geographic range and range of composition possible under that type; and a brief description of ancillary information concerning typical boundaries, associated species, and common situation. This format was designed with reference to the typical process of identification by an aerial interpreter (see figures 1 and 2). Page layout was designed to emphasize direct identification. The example color infrared stereogram and paragraph of type identification features appear at the top of the first side of each type's page. The various methods of indirect identification may be employed to confirm a choice of type, or aid the interpreter who is unable to identify a stand solely by stand characteristics. This step is often second nature to many interpreters. For this purpose, and to provide additional information about a forest cover type once it has been identified, information such as a type's geographic distribution, site preferences, typical areas of association, and associate species is also included on each key page. Where a type is identified in an unusual location, this information may even give an insight to a different soil type, bedrock, or drainage pattern that may be causing or allowing the difference.

CIR Photography. The color infrared photos presented were all chosen from those taken during the summer of 1986. Frequently a single type appears in more than one photograph in the key. This was designed to show some of the most common variations in a type, and can readily be accessed through the index. Each stereogram is labeled with location, date taken, scale and percent composition of the major component species (see figure 1).



WB/RS/BF

White Birch--Red Spruce--Balsam Fir

(Betula papyrifera, Picea rubens, Abies balsamea)

Composition: White birch, red spruce and balsam fir in various combinations constitute the major stocking

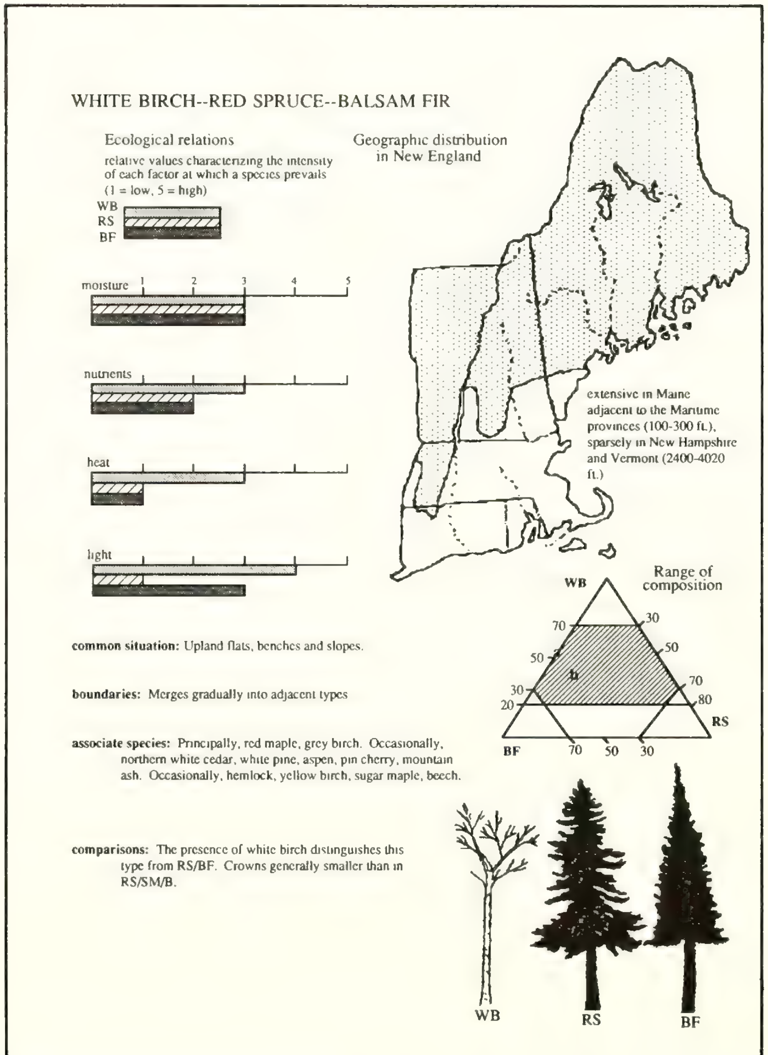
CIR
stereo pair

Cardigan State Forest, NH, 12 August 1986; a=WB 35% RS 0% BF 30%
1:6000 b=WB 25% RS 10% BF 30%

Identifying features: White Birch--Red Spruce--Balsam Fir is finely textured, as white birch is a very small-crowned hardwood. The canopy is surprisingly even, and young stands may have a carpet-like texture (refer to Fig. Y). Individual lanceolate spruce, spire-like fir and rounded birch crowns become more distinguishable in older stands. Color in CIR varies broadly, as different relative compositions have a large effect. In general, color ranges from patchy dark (more red spruce) to pink and tan (more white birch, balsam fir and other hardwood associates).

1:20000
2 September 1970

Figure 1. The front side of an example key page.



WHITE BIRCH--RED SPRUCE--BALSAM FIR

Ecological relations

relative values characterizing the intensity of each factor at which a species prevails (1 = low, 5 = high)

Geographic distribution in New England

extensive in Maine adjacent to the Maritime provinces (100-300 ft.), sparsely in New Hampshire and Vermont (2400-4020 ft.)

Range of composition

common situation: Upland flats, benches and slopes.

boundaries: Merges gradually into adjacent types

associate species: Principally, red maple, grey birch. Occasionally, northern white cedar, white pine, aspen, pin cherry, mountain ash. Occasionally, hemlock, yellow birch, sugar maple, beech.

comparisons: The presence of white birch distinguishes this type from RS/BF. Crowns generally smaller than in RS/SM/B.

Figure 2. The back side of an example key page.

Identification Features. The CIR photos were examined along with other keys and photo interpretation publications in order to draw up a list of species and stand identification features. Terms already familiar among keys and interpreters were used preferentially, but not exclusively, to describe stand pattern, texture, uniformity and color, and species crown shape and color. Although one of the advantages of color infrared film is its ability to display subtle differences in species and type color, reference to the color of types had to be treated carefully. Color infrared photography has not proven consistent enough to be able to describe a species or type by its precise hue, chroma and value. Intensity of tones and shades of color can vary considerably within a frame as well as between frames as a result of changes in sun angle, light intensity, exposure and/or variations in the processing (Enslin and Sulivan 1974, Nielson and Wightman 1971). Much information about a type can be gathered from its color, but the aerial interpreter of CIR must assess image color carefully. In addition to photography-dependent changes in color, the color observed in the image of a stand is partially dependent upon the background color of the soil, understory species or shadows, as well as the spectral reflectance of the trees. Thus, different site conditions (e.g. flooding) or stand conditions (e.g. a more uniform canopy creates fewer shadows and a less dense stand reveals more reflectance from the understory and forest floor) can shift the overall impression of color (Murtha 1971, Enslin and Sullivan 1974). Fortunately, a good quality photograph contains many clues for type identification. Here, cover types and species are described by texture, pattern, density and uniformity and relative color, both in terms of shade and intensity (see figure 1). Site and typical boundary patterns are referred to here only if they represent a principal distinguishing feature of the type (e.g. the bog in this area for Black Spruce/Tamarack). Confronting potential areas of confusion between similar types is an additional paragraph of comparisons on the reverse side of each key page (see figure 2). Similar types are mentioned, as well as how best to tell the difference between them. If a type commonly or occasionally changes with a different understory or relative species composition, it is also included in this paragraph. References to color infrared photography elsewhere in the key frequently accompany the descriptions to demonstrate some of these effects.

1:20000 Panchromatic photography. One of the drawbacks of 35 mm photography at this scale is that not much of the surrounding area, something often used by interpreters as clues to a type, is included in the photo. To make up for this deficiency, a stereogram at a smaller scale is included on the same page. Sections of the 1:20000 panchromatic ASCS photography (some taken in October, some in May of 1970 and 1974) were cut for the purpose. The 1:20000 panchromatic photography also extends the usefulness of the guide, providing an example of each type at a smaller scale and different emulsion. Each stereo pair is labeled with scale and date taken (see figure 1).

Boundaries, Common Situation and Associate Species. The reverse side of each key page contains ancillary information about each type for reference (see figure 2). As the goal is to represent the SAF types, this information about them was also taken primarily from the same 1980 publication. Consistency between key pages was important, however, so other sources were used to fill in the gaps. The pattern of boundaries observed of a particular type and some mention of the common situation it can be found in were taken from Eyre (1980) and from Silvicultural Systems for the Major Forest Types of the US (Burns 1983) and personal experience. The "common situation" may be a topographic position or special circumstances, such as "...often on sites once cleared for agriculture." The list of associate species for each type was taken directly from the list in the SAF 1980 guide.

Geographic Distribution. If the geographic position of a photo is known, some information concerning the geographic distribution of the types can be useful. Forest Cover Types (FCT) are not as well documented as species, and the SAF descriptions of the geographic distributions of the forest cover types were too vague to be of much help, so the FCT distribution was estimated to be a combination of the component species distributions, gathered from already published information (Little, 1971; Fowells, 1975). Appearing on the range map is the intersection of all the component species' ranges (see figure 2).

Ecological Relations. It is often useful to know a species' requirements and tolerances to climatic and edaphic conditions. Once a plant community has been labeled as a particular forest cover type, information regarding each species' moisture, nutrient, heat and light requirements further describes the community. Such information may also be part of the classification system. The 1954 edition of Forest Cover Types of the US and Canada, for example, lists the forest cover types according to their moisture requirements, and the 1980 edition divides the sections roughly along the lines of heat requirements (boreal, northern, central, southern and tropical). Ideally for a key to forest cover types, such information should be valid, complete and brief--i.e. reducing the information to a few words or graphic representation. As species requirements in the literature are usually presented in broad qualitative classes, while any quantitative data are scattered and necessarily refer to conditions of limited range (as species can vary greatly in their requirements over their entire range) (Lindeman 1942), other methods had to be investigated. Synecological coordinates--a method of combining the qualitative and quantitative data available into a scale of relative values--was chosen for this purpose. Developed by Bakuzis (1959) for Minnesota, synecological coordinates express, on a scale from 1 to 5, each species' requirements for "essential environmental factors . . . within a certain plant-geographical region." The term "synecology," meaning "community ecology," emphasizes the fact that the values indicate a species' environmental requirements *when competing with other plants* and not under ideal circumstances such as the absence of competition. The four essential factors used to describe species site preferences are moisture, nutrients, heat and light requirements. In general, synecological coordinates are established for a particular area by evaluating previously published information and adjusting it to the local geographic region on the basis of field observations of community (stand) species composition. The plot data recorded during the field analysis for the key supplied the community composition information used for adjusting from the Minnesota values to New England values (see figure 2).

Range of Composition Diagrams. A cover type definition includes a certain range of species composition, and to properly apply the type classification, an interpreter should know the limits of this range. The example photos in a photographic key can represent only a fraction of the variation. To give the interpreter some perspective, a diagram has been included for each type in the guide to describe the variation in species composition possible (see figure 2). To avoid confusion, the diagrams were made as mutually exclusive and exhaustive as possible, without excessively stretching any definitions in the SAF guide.

Accuracy Test

The key is intended as a guide for the interpretation of forest cover types on color infrared aerial photograpy of New England at scales of 1:6000 to 1:16000. However, any key to airphoto interpretation is only as good as the increase in accuracy one gains by its use. A simple test of accuracy was achieved for this key by comparing the results from five interpreters, each identifying 98 sample stands first without the key and then using the key, and comparing the results of their identification to the ground-truth.

The test was intended to determine (a) what levels of stand classification accuracy airphoto interpreters were likely to achieve with and without the key, and (b) the significance of any difference. CIR photography within the scale intended for use with the key was available from three sources: 1:6000 2 1/4 inch x 3 1/2 inch prints taken during the missions flown for the key in 1986, 1:8000 9 inch transparencies taken by the USFS in 1984, and 1:15840 9 inch prints taken for SPNHF in 1979. A set of 98 CIR stereo exhibits was chosen from this pool of photography. The sample photos used in the test represented all the forest cover types recognized in the key, the full range of scales intended for use with the key, and a large portion of the range of colors possible with CIR. A stand was delineated for identification on each stereo pair. Accompanying each stereo pair was a copy of a U.S. Geological Survey topographic map marked to show the location of the stand in question. Five interpreters were selected, and each was asked to go through the 98 stereo exhibits twice, first without using the key and then with the aid of the key. All the interpreters had basic familiarity with stereoscopic interpretation. Two of the interpreters were from the USFS Forest Science Lab in Durham and three were graduate students at the University of New Hampshire.

The number of samples required was calculated using a combination of acceptance sampling and the binomial probablility density function (Ginevan, 1979), and a desired classification accuracy of 85% with 95% confidence. Using both statistical procedures takes care of two sources of error, the probability of rejecting a map of high accuracy (producer's risk) and the probability of accepting a map of low accuracy (consumer's risk). The coefficient used to measure the degree of agreement between an interpreter's classifications and the ground-truth is Cohen's k, in which k=the proportion of correct responses after chance has been removed (Cohen, 1960). Testing for the significance in the difference in accuracy with and without the key, the test for significance between two independent k values was employed.

RESULTS AND DISCUSSION

Results of the accuracy test are summarized in table 1, which shows the results of all the interpreters, combined and individually, with and without use of the key. In both phases of the test, the number of errors substantially exceeded the limit of 8 imposed in the selection of sample size, and the hypothesis of 85 percent accuracy is accordingly rejected.

Without the key, all interpreters together achieved a classification accuracy of 23 percent. Deflated for chance agreement, this figure is represented by a k of .189. With the key, all-interpreter classification accuracy rose to 30 percent, represented by a k of .269. These two k's are significantly different at the .01 level, accepting the hypothesis that the key improves interpreter accuracy. Considering the interpreters individually, each was able to increase the proportion of correct identifications, and there were instances of major increases in accuracy in particular type groups (table 2). Three of the five interpreters, however, did not increase their overall accuracy significantly. One interpreter appeared to improve considerably in accuracy, but because of the high standard error associated with her results, the improvement did not turn out to be statistically significant.

Patterns in Identification. The results from each interpreter at each phase of the test were displayed in cross-classification matrices. Several patterns emerged from the study of these and the aggregate tables. The pure pine types, primarily WP and RP were the most often classified correctly. The plantation examples of RP and the distinct crown of WP probably facilitated correct classification. The hardwood types appeared to be the most difficult to distinguish. Without the key, most of the hardwood stands were identified as either WB, SM, RM, RO or SM/B/YB, and there was no distinction between RM and SM. With the key, the hardwood categories appeared much more clearly distinguished in the matrix, although SM/B and WB still seemed to pose difficulties. Surprisingly, the accuracy of the spruce-fir types was also relatively low, about as low as the accuracy of interpretation for the hardwood types.

With 21 different forest cover types as choices, many of the interpreter's fell back into a pattern of responses unique to their own way of working. One prominent pattern, revealed by the distribution of the marginals, was the tendency for some to depend heavily on the multi-species type for classification, a pattern differing distinctly from the distribution of the ground truth marginals. This pattern of distribution in the responses (and correspondingly in the marginals) sets limits on the proportion of correct responses possible, setting an upper limit on k that is less than 1, or perfect agreement. That maximum value of k set by the marginal distributions (k_M) were low, between .438 and .710 (see table 1). Thus, a substantial amount of the disagreement in the accuracy test could be attributed to marginal discrepancies. Although it is still a form of disagreement, this value can serve as an indicator of fuzzy category boundaries (Cohen 1960). Further training of the interpreters would familiarize them with the types, sharpen their understanding of those boundaries and increase accuracy.

Content, Organization and Ease of Use. Another indication of the success of the key was accessed through a brief questionaire completed by each interpreter after they had completed the test. The interpreters found the variation in color between the key photos and the test photograpy to be the biggest difficulty in interpretation, although one interpreter commented how helpful the reference photography and example photography was for comparing types and variations in types on a single photo and getting "calibrated" to the color combinations in the sample stands. One suggestion, to add a Balsam Fir type, was later incorporated into the key. Most of the interpreters found the organization of the key easy to work with, and felt that with more practice using the key, their accuracy could improve. Each interpreter, however, seemed to concentrate on certain sections of the key for their interpretation. Complete reliance on the color infrared photos in the guide would increase the difficulty of interpreting photography of other color variations and other scales. As not enough samples were run to further subdivide the data, no assessment was made of the accuracy achieved at each of the three scales.

Limitations Imposed by the Accuracy Test. Befort and Ullimann (1977) reported accuracies of 75% for the identification of habitat types in west central Idaho. The very large scale used and the simplicity of the types in the western United States account for this very large difference in test accuracy between their

Table 1. The increase in accuracy achieved using the key; as observed using the Coefficient of Agreement (k).

	Coefficient of Agreement (k)		Maximum Agreement permitted by the marginals (km)	
Interpreter	w/out key phase	w/ key phase	w/out key phase	w/ key phase
all interpreters	.189	.269	.608	.710
SC	.196	.298	.512	.578
JK	.170	.351	.540	.670
MM	.169	.180	.466	.674
BF	.244	.390	.715	.759
KS	.183	.211	.438	.581

Table 2. The proportion of correct responses, of all the interpreters combined, in each of 3 type groups: pine, spruce-fir-cedar, and hardwood.

type group	without key	with key
pines	38.1	49.0
spruce-fir-cedar	14.3	27.9
hardwoods	17.8	21.6

guide and this one designed for New England forest cover types at much smaller scales. In addition, several identifiable limitations on accuracy were imposed by the design of the accuracy test. The test included a wide range of scales (1:6000, 1:8000 and 1:15840) and three types of photography, both transparencies and prints and 9 inch and 35mm formats, with an accompanying range of color schemes and resolutions. This already imposed conditions in the test that would not be present under ordinary working conditions. In addition, no interpreter was familiar with the entire area covered by the photography, and the sample stands in the test were considered in isolation from one another. The limited experience of the interpreters restricted the accuracy achievable on a single run of the test. Only two of the interpreters had worked with color infrared before, and none were familiar with more than one of the three scales and photo types. An hour spent familarizing each interpreter with the key and the three types of photography used before the test would almost certainly have increased interpretation accuracy.

CONCLUSIONS

Using this key, photo interpreters were able to significantly increase their interpretation accuracy of New England forest cover types on color infrared photography. An increase in the "close" responses, within each type group of pine, spruce-fir-cedar or hardwood also occurred, indicating a learning process initiated by the key. The organization of the types and of the information within the key pages was easy for interpreters to use and understand. More information would have been too cumbersome. The color variation that accompanies all color infrared photography was a difficulty in interpretation, but the inclusion of additional reference photography to the key and the representation of more than one type on a photograph allowed comparison and assisted the interpreter in "calibrating" to the different photography. As pointed out by the results of the accuracy test, accurate airphoto interpretation requires experience. The success of the key was in increasing the speed at which an interpreter could gain experience, and in the short term, effectively increasing an interpreter's current level of accuracy.

REFERENCES

Bakuzis, E.V. 1959. Synecological coordinates in forest classification and in reproduction studies. PhD dissertation. Univ. of Minnesota. St. Paul. 244 p.

Bakuzis, E.V. and V. Kurmis. 1978. Provisional list of synecological coordinates and selected ecographs of forest and other plant species in Minnesota. St. Paul. 31 p.

Befort, W. and J.J. Ullimann. 1977. Airphoto guide to forest cover types in west central Idaho. PhD. dissertation. Univ. of Idaho. 140 p.

Burns, R.M. 1983. Silvicultural systems for the major forest types of the US. USDA For. Serv. Ag. Handbook no. 445. Washington D.C. 191 p.

Cohen, J. 1960. A coefficient of agreement for nominal scales. Edu. and Psych. Meas. 20(1): 37-46.

Congalton, R.G., R.G. Oderwald and R.A. Mead. 1983. Assessing Landsat classification accuracy using discrete multivarate analysis statistical techniques. Photogramm. Eng. 49(2):1671-1678.

Enslin, W.R. and M.C. Sullivan. 1974. The use of color infrared photography for wetlands assessment. pp. 697-719. In: Remote Sensing of Earth Resources, vol. III. F. Shahrokhi. (ed). Univ. of Tennessee, Tullahoma, TN.

Eyre, F.H. (ed). 1980. Forest cover types of the US and Canada. Society of American Foresters. Washington, D.C. 148 p.

Fowells, H.A. (comp.) 1975. Silvics of forest trees of the US. USDA For. Serv. Ag. Handbook no. 271. 762 p.

Ginevan, M.E. 1979. Testing land-use map accuracy: another look. Photogramm. Eng. 45:1371-1377.

Knipling, W.B. 1970. Physical and physiological basis for the reflectance of visible and near infra-red radiation from vegetation. Remote Sensing of Environment. 1:155-159.

Lindeman, R.L. 1942. The trophic-dynamic aspect of ecology. Ecology. 23:399-418.

Little, E.L. Jr. 1971. Atlas of US Trees. vol. 1. Conifers and important hardwoods. USDA For. Serv. Misc. Publication no. 1146. Washingon, D.C. US Gov't Printing Office.

Murtha, P.A. 1971. A key to air photo interpretation of forest damage in Canada. pp. 167-180. In: Color Aerial Photography in the Plant Sciences and Related Fields, Proc. Third Biennial Workshop. American Soc. of Photogramm., Falls Church, VA.

Nielson, U. and J.M. Wightman. 1971. Ultra-small-scale aerial photography forest classifications. pp. 181-193. In: Color Aerial Photography in the Plant Sciences and Related Fields, Proc. Third Biennial Workshop. American Soc. of Photogramm., Falls Church, VA.

Society of American Foresters. 1954. Forest cover types of North America (exclusive of Mexico). Washington, D.C. 67 p.

Photointerpretation Accuracy Across Two Decades of Forest Inventory

G. Deegan and W. A. Befort
Resource Assessment Unit
Forestry Divison, Minnesota Dept. of Natural Resources
2002 Airport Road, Grand Rapids, Minnesota 55744

ABSTRACT

Photointerpreted cover type identifications used in statewide forest inventories in Minnesota in 1977 and 1988 were 59 and 70 per cent correct respectively. Adjusted for chance agreement, these figures decline to 38 and 59 per cent. Although double sampling is used to adjust type acreages, low photointerpretation accuracy has an impact on sampling error of the completed inventory.

INTRODUCTION

Aerial photo interpretation ordinarily serves as the first sampling phase in large-area forest inventories, such as the state-by-state Forest Inventory and Analysis (FIA) surveys which the U.S. Forest Service conducts. In the early 1970s the state of Minnesota undertook to cooperate with the North Central Forest Experiment Station in carrying out the Minnesota survey. Data from the first cooperative effort were published with the nominal date of 1977, and a second cycle of cooperative statewide inventory is now under way, with a nominal date of 1988. The 1988 survey, the first in Minnesota in which identical ground plots from a previous survey were remeasured, is certain to provide a wealth of comparative data about forest conditions and trends between the 1970s and now. Although as yet only partial and preliminary data are available for 1988, the two surveys, with their nearly identical methodologies, also afford opportunities for comparative study of the effectiveness of data-gathering techniques such as photointerpretation.

AERIAL AND GROUND DATA

Like all such statewide inventories, the Minnesota surveys were two-phased in structure, with initial photo interpretation of about 275,000 systematically spaced plots followed by detailed ground measurement of a subsample of about 9,000 plots. The numbers of photointerpreted points in each cover type and condition class provide factors by which the ground data may be extrapolated to total acreages and volumes across the state. Numerous laboratory studies have addressed questions of forest cover type identification accuracy in photointerpretation and other kinds of image analysis; we thought it might be interesting to find out what kind of result was being achieved in operational practice with available aerial photography. The inventory design lends itself well to such an inquiry: cover type identity of each ground-measured forest plot is determined by strict procedural rules, and agreement with ground-determined cover type is the sole criterion of photointerpretation accuracy. Here the much-abused term "ground truth" can actually be employed with some rigor.

At this point in the 1988 inventory cycle (mid-1989), ground truth data are available only from portions of the five-county northeastern Minnesota inventory unit, the so-called "Aspen-Birch Unit". We assembled all the currently accessible 1989 data from commercial forest plots in this unit--1,742 field-measured plots--and constructed a photo-vs.-ground cover type error matrix from it, then did the same with the 3,275 ground-measured commercial forest plots from the 1977 inventory of the same unit. The results appear in Tables 1 and 2. It should be borne in mind that the 1988 data are preliminary, incomplete and unedited.

In any such interpretation exercise, and especially where one cover type is disproportionately represented on the ground as is our aspen-birch type, human or machine interpreters inevitably get some of their calls right by mere chance. It is common practice to adjust the percentage of agreement in an error matrix by calculating the KHAT statistic (Cohen 1960, Congalton and Mead 1983), a deflated percentage-correct figure with the "lucky guess" factor removed. KHAT has the further advantage of permitting statistical comparisons between independent error matrices. We worked out KHAT and its standard error for each matrix. The difference between the two KHATs is significant at the .0001 level.

Table 1. Photo vs. ground cover type, 1977: numbers of inventory plots.

	Ground										
Photo	JP	R/W	S-F	BS	WC	TK	LH	NH	A-B	N/O	TOTAL
Jack Pine	**30**	6	9	8					12		65
Red/White Pine	3	**6**				1			7		17
W.Spruce-B.Fir	7	5	**150**	83	67	10	27	3	86	3	441
Black Spruce	6	1	28	**246**	33	56	1		18	7	396
White Cedar				10	**17**		2		1	3	33
Tamarack	1	1	3	14	12	**21**				1	53
Lowland Hdwd	1		21	21	33	13	**22**	6	42	5	164
North. Hdwds			5		3		6	**28**	22		64
Aspen-Birch	11	36	191	65	38	10	117	105	**1416**	18	2007
Nonstock/Other	2		4	6	1	4			18		35
TOTAL	61	55	411	453	204	115	175	142	1622	37	3275
No. Correct	30	6	150	246	17	21	22	28	1416	0	1936
% Correct	.492	.109	.365	.543	.083	.183	.126	.197	.873	.000	.591
									KHAT		.379
									Stderr KHAT		.013

Table 2. Photo vs. ground cover type, 1988: numbers of inventory plots.

	Ground										
Photo	JP	R/W	S-F	BS	WC	TK	LH	NH	A-B	N/O	TOTAL
Jack Pine	**8**	2	2	5					8		25
Red/White Pine	2	**17**	12	1	1			2	15		50
W.Spruce-B.Fir	7	1	**75**	14	15	4	3	7	48		174
Black Spruce	1		17	**241**	20	17	3		14		313
White Cedar			11	7	**78**	2	2		7		107
Tamarack			3	12	6	**55**	1		1		78
Lowland Hdwd			6	2	4	2	**63**	2	30		109
North. Hdwds			3		1		2	**30**	10		46
Aspen-Birch	5	12	66	13	13	4	45	30	**652**		840
Nonstock/Other											0
TOTAL	23	32	195	295	138	84	119	71	785	0	1742
No. Correct	8	17	75	241	78	55	63	30	652		1219
% Correct	.340	.531	.385	.817	.565	.655	.529	.423	.831		.700
									KHAT		.588
									Stderr KHAT		.015

It is difficult to account for the spread. The same typology and the same general photointerpretation technique were used on both occasions. One difference lay in the kind of photography employed. The 1977 interpreters used 9x9" format 1:15,840 black-and-white summer infrared photography under stereoscopes of moderate (2x-4x) magnification. The 1988 interpreters used 1:58,000 National High Altitude Photography Program color infrared imagery under high (8x) magnification. Undoubtedly the color infrared film helped the 1988 interpreters make some distinctions (especially among lowland conifers) that eluded the 1977 interpreters, but this should have been counterbalanced by the fact that the 1:58,000 photography was taken when leaves were off the deciduous trees.

Sampling Error Effects

In inventory practice, of course, photo-derived cover type acreage estimates are systematically adjusted on the basis of discrepancies discovered in the field, so that the consequences of interpretation errors are less grave than the uninitiated might fear. However, any loss of correlation between photo and ground necessarily weakens the confidence we can place in photo-based extrapolations of ground data. To illustrate this effect we calculated the expected 1977 sampling errors for each cover type in the Aspen-Birch Unit under three differing conditions: 1) no photo misclassifications, 2) actual 1977 misclassification rates, and 3) 1988 misclassification rates (Husch et al. 1972, pp. 246-248). Results are given in Table 3. As might be anticipated, there is a wide gap between the theoretically attainable sampling error and the results achieved in practice; however, the significant difference between 1977 and 1988 misclassification rates results in large sampling error differences in only a few cover types, chiefly lowland conifers.

Table 3. Effect of photo misclassification on survey error.

In 5,000,000 ac. commercial forest; 20,000 photo & 3,275 field plots.

Cover type	A. Approximate total type acreage	B. Sampling error with no misclassifications	C. Error at 1977 misclassification rate	D. Error at 1988 rate
Jack Pine	150,000	4,931	10,627	10,360
Red/White Pine	140,000	2,451	11,010	9,626
W.Spruce-B.Fir	640,000	12,069	28,105	26,763
Black Spruce	690,000	11,527	26,024	20,915
White Cedar	300,000	3,531	20,209	15,946
Tamarack	140,000	4,463	15,209	11,683
Lowland Hdwd	230,000	7,711	19,512	19,583
North. Hdwds	200,000	4,889	17,134	15,893
Aspen-Birch	2,510,000	17,222	37,799	36,318
Total Acreage	5,000,000			

The acreage adjustment process, from which the sampling errors are derived, is rather sensitive to the original pattern of distribution of photointerpretations among cover types. In the present case we assumed that 1988 plots were as numerous as 1977 plots and distributed similarly; only the error matrix was altered in the acreage adjustment. One result is that the distribution of the expected errors is somewhat skewed; thus the 1988 sampling errors for jack pine and aspen-birch are slightly smaller than the 1977 errors, even though the 1977 interpretations were better in these two types. Conversely the 1988 error in lowland hardwoods is slightly larger than the 1977 error, though 1988 interpretations in this type were more accurate. A better comparison will be possible when final 1988 data are available.

Sources of Misclassification: Type Composition

The similarities between the 1977 and 1988 error matrices are more obvious than their statistical differences. Row for row and column for column, misclassifications tend to occur in the same pattern; the distribution of empty and near-empty cells is much alike in each matrix. It is plain that the same process went on in each case and that the same kinds of mistakes tended to occur. Errors may arise from several sources. First and most obviously, the kind of aerial photography in use may not provide enough information for interpreters to distinguish tree species and other cover type characteristics with consistency. This deficiency of information is always present in some degree, and varies with scale, season and film type. A large literature addresses this question, which is not dealt with here. But a subtler source of error lies in the character of the cover types as defined by their species composition. Pure stands of any type rarely occur in the wild, and every stand is in a continual process of change. For inventory purposes, arbitrary rules have been devised which result in each plot's being assigned to a cover type; but the same type assignment can prevail over a wide variety of composition, and no one is more aware than the field inventory forester how often his final cover type determination differs from his initial impression of the plot. If the field forester may be deceived, how much more may the photointerpreter, who can see only the overstory trees--and those imperfectly--be led astray in assigning a plot to a cover type. The clearest example of this kind of error lies in the cell--well populated in both matrices--containing spruce-fir plots mistakenly assigned to the aspen -birch type. Anyone familiar with northern Minnesota forests will have no difficulty in supposing that many of these plots bore a thin overstory of mature aspen or birch and a thriving successional population of balsam fir. In this case the materials for a correct diagnosis were often simply not available to the interpreter, and would not be available from any photograph except at very large scale. We thought it might be illustrative to create a table showing the degree to which cover types tend in fact to be stocked by their nominally dominant components, and the extent to which other type components tend to participate in them. This appears as Table 4.

Table 4. Percentage composition of FIA cover types by cover type component, Aspen-Birch Unit, 1977.

Component (% growing stock volume)	**Cover Type**								
	JP	R/WP	S-F	BS	WC	TK	LH	NH	A-B
Jack Pine	**73.7**	6.3	0.9	3.7	0.7	0.4	0.2	0.9	1.2
Red/White Pine	2.7	**72.7**	3.7	1.1	2.4	0.8	0.6	1.7	2.4
W.Spruce-B.Fir	2.5	8.0	**49.2**	8.4	10.5	2.8	10.7	10.7	12.2
Black Spruce	6.3	0.9	7.4	**69.9**	10.4	16.0	0.9	0.5	1.8
White Cedar		0.3	5.9	1.6	**59.9**	3.5	5.1	5.8	1.3
Tamarack			1.0	5.7	1.5	**72.6**	0.2	0.1	0.1
Lowland Hdwd		0.1	0.9	0.4	2.4	0.2	**64.9**	11.7	3.4
North. Hdwds	0.6	0.4	0.6		0.3	0.1	2.6	**41.3**	2.2
Aspen-Birch	14.1	11.4	30.3	9.2	11.9	3.7	14.7	27.3	**75.5**

The table shows, at the very least, that aspen-birch species are frequently encountered in most cover types. Given their role as pioneer species, they may often be in a position to mask the true type identity of inventory plots as viewed on the aerial photograph. It also shows that the spruce-fir, white cedar and northern hardwood types, as sampled in inventory plots, tend to be mixed in composition, a fact with obvious implications for photointerpretation. Correlation (r) of the diagonal elements with 1988 percentages of correct classification in each type was .52, nonsignificant at the .05 level.

Sources of Misclassification: Plot vs. Stand Interpretation

The foregoing suggests that another cause of frequent photo misclassification of inventory plots may be the fact that an arbitrarily placed plot rather than a delineated forest stand is the subject of interpretation. Interpretation of isolated points is always a severe test of an image analysis technique. Working on a plot-by-plot basis, the interpreter has little opportunity to organize the contextual information the photo provides and assemble it into a satisfactory over-all solution. Moreover, the plots often fall in mixed growth at boundary or transition zones between well-defined types. The interpreter's situation is analogous to that of the analyst of digital satellite imagery who attempts a pixel-by-pixel cover type mapping: many pixels lie in transition areas with mixtures of species, and failure to classify these reliably may detract severely from results achieved in more homogeneous areas. Table 5 presents results of a different photointerpretation approach: drawn from northern Minnesota timber type mapping data, it shows the frequencies of interpretation of 216 randomly chosen delineated stands rather than isolated plots. These delineations were made on 1:15,840 black-and-white infrared photography like that used in 1977 FIA work.

Table 5. Photo vs. ground cover type: stand identification.

	Ground										
Photo	Asp	Bir	BSL	BFir	WCed	Balm	BAsh	Tck	WSpr	JPin	TOTAL
Aspen	**92**	1	1	4	2	1	3				104
Birch		**1**									1
Bl.Sp.Lowld.			**37**	1	2			10			50
Balsam Fir				**2**	2						4
White Cedar	2		2		**15**		1				20
Balm of Gilead	1										1
Black Ash	2										2
Tamarack			5		6			**18**			29
White Spruce									**2**		2
Jack Pine										**3**	3
TOTAL	97	2	45	7	27	1	4	28	2	3	**216**
% Correct	.95	.50	.82	.29	.55	.00	.00	.64	1.00	1.00	**.78**
									KHAT		**.70**
									Stderr KHAT		**.04**

The KHAT is significantly different from that of Table 2 at the .01 level. While scarcely conclusive owing to the small sample size and the differing set of cover types, these results from type-mapping work suggest that plotwise interpretation may result in lower accuracy than is usually associated with aerial photographic cover type delineation.

A recent University of Minnesota study involving satellite cover type mapping shows the same principle at work. Moore (1987) used 30-meter Landsat Thematic Mapper data to classify 14 cover types in the Itasca State Park area. The nine forest types identified were similar to those in Tables 1-4 above. She recorded two measures of classification performance: per-pixel performance, in which all pixels in the scene were evaluated for accuracy, and test-field classification, in which mixed-pixel areas were filtered from the scene and accuracy assessment was restricted to the more homogeneous portions. KHAT in the per-pixel evaluation was .304; KHAT in the test-field evaluation was .503, a difference significant at the .0001 level.

CONCLUSIONS

The plotwise forest cover type identification accuracy levels presented here are probably fairly representative of what is achieved in operational FIA work throughout the region. Significant differences in accuracy occur from one cycle to another--and perhaps from unit to unit within a given cycle as well-- but the reasons are difficult to trace. The process of acreage adjustment in double sampling removes the worst effects of interpretation error so long as there is some positive correlation between photo and ground type identification, but lower photointerpretation accuracy adversely effects the standard error of type area estimates. Low photointerpretation accuracy is correlated--albeit loosely--with high within-type species diversity, and the point-by-point approach to image interpretation seems to show much the same weaknesses in stereo photo work that it exhibits in digital image analysis. It is worth remarking that reported accuracies in a number of satellite-based type-mapping studies are not greatly different from the FIA photointerpretation accuracies discussed here.

LITERATURE CITED

Cohen, J. 1960. A coefficient of agreement for nominal scales. Edu. and Psych. Meas. 20(1):37-46.

Congalton, R.G. and R.A. Mead. 1983. A quantitative method to test for consistency and correctness in photointerpretation. Photogram. Eng. 49(1):69-74.

Husch, B., C.I. Miller and T.W. Beers. 1972. Forest Mensuration. Ronald Press, New York. 410p.

Moore, M.M. 1987. Classification of forest vegetation in north-central Minnesota using Landsat multispectral scanner and thematic mapper data. Unpubl. Ph.D. diss., U. of Minn., St. Paul. 166p.

On Using Non-Imageable Point for Measuring Object Height on Vertical Aerial Photographs

C. J. Liu
Associate Professor
Department of Forestry, University of Kentucky
Lexington, Kentucky 40546-0073 U.S.A.

ABSTRACT

A new equation using geometric relationships between the object's image and the non-imageable point on a single vertical aerial photograph was recently developed by Chinese forest researchers for measuring object height. The original paper was written in Chinese and the journal containing the paper has very limited ciruclation outside China. This paper serves as an exposition of the original idea, an interpretation of mathematical derivations, and a commentary on uses of the new object height equation. Translated also is the original application example.

INTRODUCTION

Accurate measurement of tree heights on aerial photographs is important to forest practice and research. Although many photogrammetric height determination methods are available, object heights must be measured on the ground (as ground truth) to check the accuracy of photo measurements. In forestry applications, such as aerial photo mensuration, a large number of tree heights needs to be checked on the ground. Thus, the task of tree height determination is arduous. The non-imageable point is a characteristic point unknown to many. Few if any have endeavored in a research or practice pertaining to the non-imageable point. This paper (Peng et. al 1988) introduces a procedure which uses the non-imageable point to determine the sun's elevation and then calculates the object height based on measured shadow length. Repeated experiments indicate that this new procedure, using the characteristic of the non-imageable point, is efficient in determining object heights.

GEOMETRY OF THE NON-IMAGEABLE POINT

Let T represents a flat surface in the imaged area, t the vertical photo, and α the sun's angular elevation at the time of photography. Furthermore, let S be the lense of the aerial camera, o' and O be the principal point in the photo plane t and the exposure station on the ground T, respectively. As depicted in Fig. 1, AB and NM represent two objects on the ground. SQ, BB", and MM" are parallel light rays projecting the center of the camera lense and tips of the ground objects as points Q, B" and M" in T. Line segments OQ, B'B", B'Q, M'M", and M'Q connect their respective end points in T.

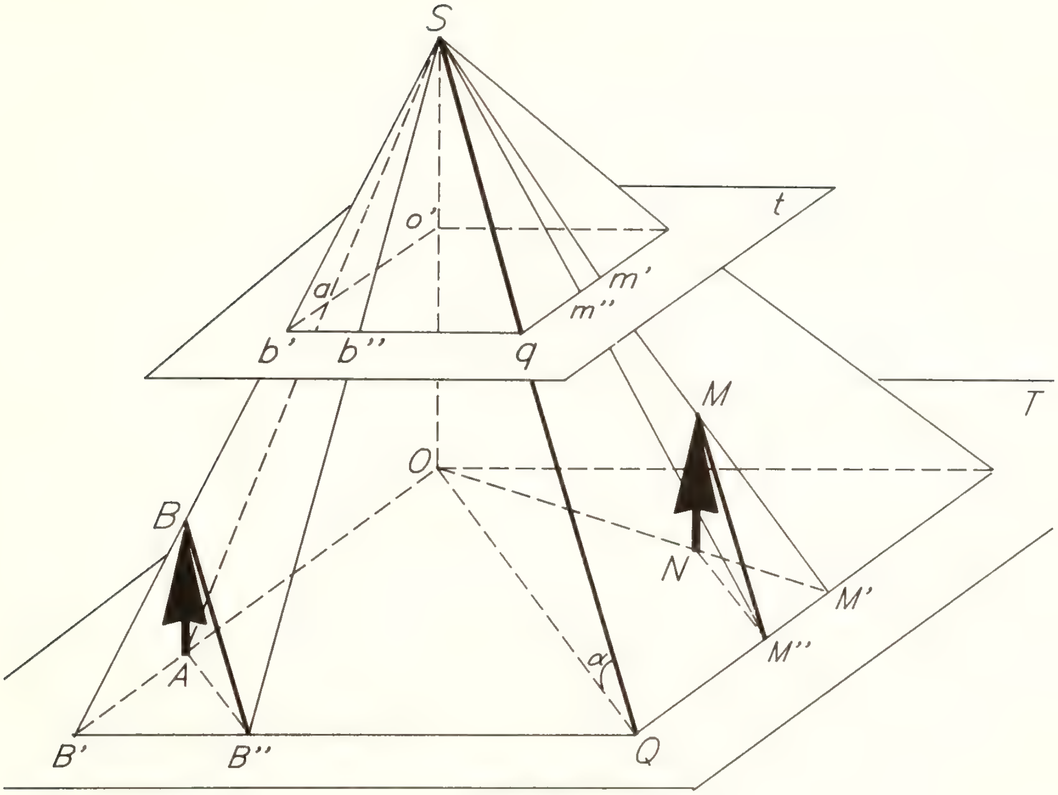


Figure 1. Location of the non-imageable point on a vertical photograph.

The non-imageable point is the point of intersection of two or more extended lines, each connecting the tip of an object and the tip of the object's shadow on a single vertical aerial photograph. The non-imageable point q in Fig. 1 is the intersection of extentions of line segments b'b'' and m'm''. To utilize the non-imageable point for object height determination, a new height equation must be developed. Prior to the derivation of the height equation, we show that all lines connecting the tip of an object and its shadow meet at one point on a vertical aerial photo and that this intersection point is located at a point that is the projection of the camera.

Consider first that the ground in the photographed area is level. Since $SO \perp T$ and $BA \perp T$, we write $\angle$ SQO = $\angle$ BB''A = α. It follows the Rt$\triangle$BAB'' ~ Rt$\triangle$SOQ, and accordingly, AB'':OQ = BA:SO.

Because points S, B, and B' are collinear and BA//SO, points B', A, and O must be collinear. Thus, $\angle$ B'BA = $\angle$ B'SO and $\triangle$BAB'~$\triangle$SOB'. We then have BA:SO = B'A:B'O. Combining this result and the one derived in the preceeding paragraph, we yield AB'':OQ = B'A:B'O.

Because SO//BA and BB''//SQ, plane BAB'' and plane SOQ are parallel. Then, in Rt$\triangle$BAB'' and Rt$\triangle$SOQ, AB''//OQ. This leads to $\angle$ B'AB'' = $\angle$ B'OQ which, together with the result yielded at the end of the proceeding paragraph, implies $\triangle$AB'B'' ~ $\triangle$OB'Q. We now have the congruence of angles, $\angle$ AB'B'' = $\angle$ OB'Q, which indicates the B'B'' and B'Q are collinear.

Similarly, it can be shown the M', M", and Q are collinear. We may now state that line segments B'B" and M'M" intersect at point Q-the position of the shadow casted by the camera (or the aircraft). Based on projection theory, two lines must intersect on a vertical photo if their corresponding lines intersect on the ground. Therefore, the corresponding intersection point in t is the non-imageable point q.

Next, consider the case when the ground is not level. Since BB"//SQ, therefore BB" and SQ determine the plane SBB"Q. Relationships $B' \in SB$ and $SB \in SBB''Q$ implies $B' \in SBB''Q$. In addition, $B'' \in SBB''Q$ inplies $B'B'' \in SBB''Q$. For both B'B" and SQ in SBB"Q, we may intuitively think that $B'B'' \cap SQ$ at point Q' which differs from Q because the ground is not level.

Because aerial photography uses the method of central projection, points b', b" and q', which are projected images of B', B", and Q' in SB'B"Q, must be on a line that is the intersection of plane SB'B"Q and t. This is equivalent to saying that b', b", q' are collinear points in t.

For the same reason, we say that $M'M'' \cap SQ$ at Q" and that a line connecting m', m", and q" must be the intersection line of SMM"Q and t. Since SQ and t intersect at only one point, q' and q" converge at a single point q in t. We have proved that lines connecting tips of objects and tips of objects' shadows must meet at only one point—the non-imageable point.

A HEIGHT EQUATION BASED ON THE NON-IMAGEABLE POINT

Let plane t be a single vertical aerial photograph and T' the plane containing line segment BB" which is the shadow of tree AB. Angle β (<0) between T' and a hypothetical horizontal plane T is a measure of slope at the tree's base A (Fig. 2). As shown in Fig. 2, SQ represents the light ray passing through the camera lense S and intersecting the plane T at Q. The point O is the vertical projection of S onto plane T. Note that the ground exposure station and the non-imagable point on the ground may or may not be in T. Line segment $ac \in t$ is the image of the tree's shadow $AB'' \in T'$. Points a, c, b, o', and q in t correspond to point A, C, D, O and Q in T.

Draw a perpendicular from B" ($\in T'$) to T and designate the point of intersection D. Since $B''D \perp T$ and $SO \perp T$, thus B"D//SO. Accordingly, B"D and SO determine the plane SOB"D and $\angle DAB'' = \beta$.

Connect SB", AD and OD so that SB" intersect OD at point C in T and then let $\angle DOQ = \delta$. Since $AC \in T$ corresponds to $ac \in t$, we immediately have $AC = m \cdot ac$ in which m is the point scale at the base A. Let $ac = \ell$, we write $AC = m \cdot \ell$.

Let $\angle OSB'' = \gamma$ and $\angle SQO = \alpha$. It follows that $\angle CB''D = \gamma$ Since B"D//SO. Because the sun's elevation is the same everywhere in the photographed area, therefore, BB"//SQ. By definition, we have AB//SO. Thus, SOQ//BAB"D. When a plane intersects two parallel planes, the two intersecting lines must be parallel in the intersecting plane (in this case plane T), i.e., AD//OQ. It follows $\angle ADO = \delta$.

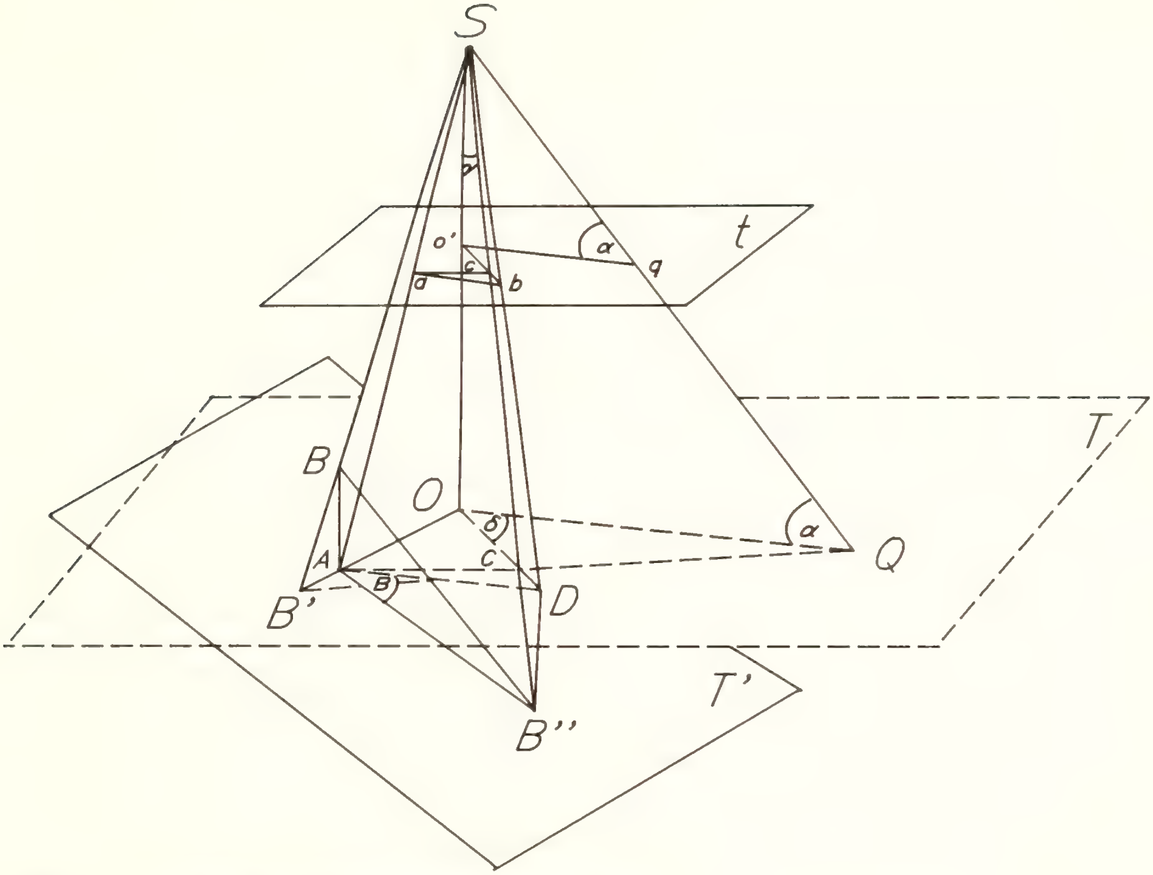


Figure 2. Relationships between an object and its shadow.

Since AB//SO and B"D//SO, we have AB//B"D. It follows that points A, B, B", and D are coplanar and $\angle$ AB"B = $\angle \alpha - \angle$ DAB" = $\alpha - \beta$.

Let tree height AB = h and employ the law of sine. It follows that, in $\triangle$ABB", AB:sin$\angle$ AB"B = AB":sin$\angle$ ABB". After substitution, we have AB" = $h \cos\alpha / \sin(\alpha - \beta)$.

We also know that, in Rt$\triangle$AB"D, AD = AB" $\cos\beta$ = $h\cos\alpha\cos\beta/\sin(\alpha-\beta)$ and B"D = AB"$\sin\beta$ = $h\cos\alpha\sin\beta/\sin(\alpha-\beta)$. In addition, in Rt$\triangle$CB"D, CD = B"D $\tan\gamma$ = $h\cos\alpha\sin\beta\tan\gamma/\sin(\alpha-\beta)$.

According to the law of cosine, in Rt$\triangle$ACD, $AC^2 = AD^2 + DC^2 - 2AD \cdot DC\cos\delta$ which implies

$$(\ell \cdot m)^2 = \frac{h^2\cos^2\alpha}{\sin^2(\alpha-\beta)} (\cos^2\beta + \sin^2\beta \cdot \tan^2\gamma - \sin 2\beta \cdot \tan\gamma \cdot \cos\delta)$$

or,

$$h = \frac{\sin(\alpha-\beta) \cdot \ell \cdot m}{\cos\alpha \sqrt{(\cos^2\beta + \sin^2\beta \cdot \tan^2\gamma - \sin 2\beta \cdot \tan\gamma \cdot \cos\delta)}} \tag{1}$$

Which is the equation for calculating object height when $\beta < O$.

When $\beta = O$, i.e., the ground is level, object height is calculated by $h = \tan\alpha \cdot \ell \cdot m$ (2)

When $\beta > O$, we may simply change the sign of β. Eq. 1 becomes

$$h = \frac{\sin(\alpha + \beta) \cdot \ell \cdot m}{\cos\alpha \sqrt{(\cos^2\beta + \sin^2\beta \cdot \tan^2\gamma + \sin 2\beta \cdot \tan\gamma \cdot \cos\delta)}} \qquad (3)$$

To obtain object height, the first step is to locate the non-imageable point–the point in a single photograph that is the intersection of two or more lines each connecting the image of the tip of an object and its corresponding shadow images (the tip of the object's shadow). The second step involves measuring or calculating 4 angles and 2 line segments. The line $\ell = ac$ can be measured on the photo and m is the point scale at a that may be calculated with reference to a topographic map. The sun's angular elevation, assumed to be the same for all photos in a flight line, can be determined by the equation

$$\alpha = \tan^{-1}(f/o'q) = \tan^{-1}(f/d) \qquad (4)$$

in which the camera focal length $f \cong So'$, and $d = o'q$ is measurable on the photo t. Angle β can be determined from viewing the stereopair under the stereoscope. The angle δ between o'q and o'c is measurable. Angle γ can be obtained by substituting known quantities into the equation $\gamma = \tan^{-1}(o'c/f)$.

EXPERIMENTAL RESULTS

Many experiments have been conducted to verify the accuracy of this newly developed height equation while results from one such experiment are reported here. The study area is located at North 27°51' latitude and East 113°8' longitude in a hilly area near the Zhu-Zhou District, Hu-Nan Province, People's Republic of China. Aerial photos used (80530933–80530916) were taken at a flying height of 940 m above sea level. The focal length of the camera is 114 mm while the average photo scale is 8,000. In addition, 10,000 scale topographic maps, mirror and pocket stereoscopes, parallax bar, engineering rule, transit and leveling rod were used in field or office.

Two methods can be used to determine the position of the non-imageable point q and subsequently the line length, $d = o'q$. The first method graphically determines the position of q and is most accurate if object image and shadow are clearly visible on the photograph. By viewing aligned stereopair under the stereoscope, one draws at least 3 lines connecting tips of trees and corresponding tips of trees' shadows. Two such lines, preferably perpendicular to each other, are used to determine the non-imageable point. If the position of q is located improperly, the three lines will form a triangle around the theoretical non-imageable point after the third line is drawn. Experience indicates that none of the sides of such a triangle should be longer then 1 mm, to obtain accuracy in the calculation of object height. Since images of trees and their corresponding shadows are not large (several minimeters or shorter) on 8,000 scale photographs, the error is usually greater (occasionally longer than several minimeters) than what is acceptable. This is particularly true when points to be measured are close to each other. To insure the accuracy required in determining the position of the non-imageable point, several trials are often necessary.

The second method obtains d numerically by employing the following geometric relationship

$$h = \tan\alpha \cdot m \cdot \ell = (f/d) \cdot m \cdot \ell$$

which gives

$$d = (f/h) \cdot m \cdot \ell$$

This method required field measurement of the height h of trees that is ideally near the exposure station and is not on rugged terrain. Using this method, differences among d's of objects within the same photo usually are less than 1 mm. Compared with the graphic method, the quantity d determined by the numerical method is more accurate, by a factor of 10. For each photo, only one object height needs to be measured in the field.

Since the sun rises at a constant speed and the speed of the aircraft is also constant, lenghts of d (o'q) on photos in the same flight line can be calculated by interpolations from d's of the two end photographs.

Shadow length measured by straight rule has a large variance and thus less precise. Experience indicates that more accurate readings can be obtained by using a parallax bar to measure shadow length under stereoscopic viewing (mainly for magnification). In our practice, we used the mean of 3 repeated measures of the same shadow length and required that the maximum difference among 3 measured shadow lengths be no greater than 0.05 mm.

Since the experimental area was relatively flat, we used Eq. 2 to calculate object height. The point scale at the base of the object was calculated by the equation $m = (H-E)/f$, in which H is the flying height above mean sea level and E is the elevation at the base of the object.

To learn about the accuracy associated with the new height equation, we calculated object heights both in the office and in the field. Results are shown in Table 1. The error ranged from 0.21 to 27.95% and the mean error was 8.269%. In other words, the average accuracy was 91.731% which was higher than required by field measurement with surveying instruments. Among all readings, 81% of the measurements attained accuracies higher than 85%. Three unusally large readings within the data were analyzed and attributed to inacurate shadow-length measurements due to obstructed shadows.

ANALYSIS AND CONCLUSION

For accurate height measurement with the new height equation, care must be taken to insure that (1) large scale, recent photographs are used, for correction for tree height growth must be exercise if photographs were taken long ago; (2) accurate readings of d is necessary, for inaccurate measurement of d resulting in systematic overestimation or underestimation; (3) objects with distinct shadow without obstructions at the base shall be chosen for measurement; and (4) shadows near the center of the photograph are used to reduce errors in the photographs.

Table 1. Accuracy of the new height equation.

Object Number	ℓ	m	d	tanα	h	h'	ζ	$\lvert\zeta$-100$\rvert$
1	1.03	7.85	105.00	1.0857	8.78	8.67	101.27	1.27
2	0.89	7.85	105.00	1.0857	7.59	7.10	106.90	6.90
3	0.96	7.85	105.00	1.0857	8.18	7.70	106.23	6.23
-	----	----	----	----	----	----	----	----
35	0.53	7.84	105.80	1.0775	4.48	3.90	114.87	14.87
36	1.40	7.84	105.80	1.0775	11.82	11.40	103.68	3.68
37	0.66	7.85	106.05	1.0750	5.57	5.70	97.72	2.28
Mean							10.216	8.269

Note: f = 144mm; $\tan\alpha = f/d$; $h = \ell \cdot m \cdot \tan\alpha$ in meters; h' = object heights in meters measured by transit; $\zeta = h/h'*100$; $\lvert \zeta - 100 \rvert$ is the accuracy.

Conclusions reached in this investigation are: first, the new equation is easy to learn and apply, saves time and increases efficiency; second, the accuracy obtainable from this technique is 91.7% which is higher than that obtained from ocular measurement of tree height in the field; third, the new technique is unsuitable for tree height measurement in dense forest stands where overabundance of shadows and obstructed tree bases prevent accurate measurement of shadow length, due to the inability of the observer to see the ground at the base of the tree.

COMMENTS

The revelation of the non-imageable point is novel and the proofs furnished in the original paper (Peng et. al 1988) are correct and concise. Traditionally, the sun's elevation is calculated by a complex equation requiring the use of special astronomical tables (i.e., a solar ephemeris) and information on time of photograph, month and day of photography, lattitude and longitude of photography, and exact scale of photography. In contrast, the new equation for sun's angular elevation requires locating the non-imageable point on the photograph and the knowledge of camera's focal length, a quantity readily available. One significant contribution from this work is the simplification in the determination of sun's angular elevation, at least theoretically.

The new height equation has been derived directly and correctly, although much insight on the interrelationship among various projected line segments is required. On rough terrain, the complex equations (Eq. 1 or Eq. 3) requiring determinations of four angles and two lengths is not simple to use. More importantly, the measurement error of one quantity will compound with that of others, slight errors in judgement affect tree height readings to a greater degree.

This new technique is essentially another shadow method of object height determination. It shares advantages and disadvantages associated with the more conventional shadow-length method. Shadows obstructed by trees or other obstructions cannot be measured. On rough terrain the observer must be able to evaluate the probable ground level at the base of the tree being measured. Accurate shadow measurements are possible only when distinct and unobstructed shadows can be found on highly contrasting surfaces.

On photographs taken near the middle of the day, the shadows are frequently too short to be measured with precision. In this as well as the more conventional shadow-length method, only individual tree heights can be measured. The accuracy of a stand height value, therefore, depends upon the judgement of the observer in estimating the relation of stand height to heights of trees measured.

The length of the shadow is influenced by the slope on which it falls. The correct height of the tree can be calculated directly from the length of its shadow providing the angle of slope is known. The precision of shadow-length measurements improves with larger-scale photographs because of more accurate shadow measurements. For accurate results, shadows should be carefully measured (to the nearest 0.1 mm or less) by special micrometer devices and magnifying monoculars.

REFERENCE

Peng, Pinxun, G.M. Tong, and M.H. Liu. 1988. [Estimate of tree height on dotless air photos.] Forest Science and Technology 215(4):17-20. Published by the Researach Institute of Scientific and Technological Information, Chinese Academy of Forest Science, Beijing, PRC.

Organizing Statewide 35mm Photography in Support of Forest Inventory

W.A. Befort and G. Deegan
Resource Assessment Unit
Forestry Division, Minnesota Dept. of Natural Resources
2002 Airport Road, Grand Rapids MN 55744

ABSTRACT

Pilot studies in the early 1980s indicated that significant economies might be achieved by using 35mm aerial photography to reduce the number of field plots remeasured in forest inventory. These studies prompted the Minnesota Department of Natural Resources to undertake a statewide 35mm photographic project in support of the current Forest Inventory and Analysis cycle. Over 9,000 commercial forest plots were eventually photographed, and continuous stereo coverage of more than 27,000 square miles was obtained. This paper describes technical and organizational features of the 35mm project, compares experienced costs against those suggested by early research, and estimates savings effected in the inventory.

INTRODUCTION

Small-format aerial photography taken by the using agency rather than by a specialized firm--the kind often dignified by the abbreviation "SAP" for "supplementary aerial photography"--is now so widely practiced and its techniques are so well documented that additions to the literature are almost redundant. But we thought it might be of interest to report a recent application in Minnesota which was unusually large in scope for a 35mm stereophotography project.

Most workers in this field owe a large debt to Prof. M.P. Meyer of the University of Minnesota, who in a long series of articles and research reports from the College of Forestry's Remote Sensing Laboratory demonstrated the techniques now in use and the economics justifying them. His research has encompassed a broad range of applications at varying scales; the particular domain of concern here is that of large-area 35mm stereoscopic coverage at medium scale. Stereo photography of county-size areas was once generally assumed to be the exclusive province of specialist firms. Meyer's work (e.g. Meyer et al. 1982) showed it to be within the reach of generalist foresters using over-the-counter 35mm equipment and rental aircraft, and demonstrated that 1:35,000 color infrared stereo print coverage could be generated for under $3 per square mile. Collaterally, Meyer and others investigated the potential application of supplementary photography to the nationwide forest surveys conducted by the U.S. Forest Service's regional forest experiment stations. A project conducted in cooperation with the North Central Forest Experiment Station (Meyer et al. 1981, Hackett et al. 1983) showed that forest disturbance could be effectively detected on Forest Inventory and Analysis (FIA) permanent plots using 1:15,840 normal-color prints enlarged from 35mm negatives, at a cost under $25 per plot. At the same time it was shown (Hahn et al. 1981) that the efficiency of FIA survey could be improved by applying a computer-based growth simulation model to previously measured but undisturbed plots, reserving expensive ($250 per plot) field examination for those plots which 35mm photography showed to have been disturbed since the previous inventory cycle.

In the current FIA inventory cycle in Minnesota these photographic and computer techniques were put to operational use. The Resource Assessment Unit of the Forestry Division, Minnesota Department of Natural Resources, planned and flew continuous normal-color stereo photography of most of the forested counties in northern Minnesota, interpreted the photographs to detect disturbance, and effected a substantial reduction in the number of ground-remeasured plots. The resulting photographs have been in high demand for other purposes and a substantial "aftermarket" has developed.

METHODS

The primary objective of the photo project was to obtain 1:15,840 print coverage of the 7500-odd commercial forest plots in the two northern Minnesota FIA survey units for which the state had undertaken inventory responsibility. Initial plans to obtain plotwise coverage in the manner of the Hackett study were dropped because plots in these units were so thickly distributed that continuous full-county coverage would be no more expensive. Normal-color film was selected as answering the basic informational requirements of the project and as being cheaper and less demanding than color infrared. Originally the end product was to be in the form of 4x6" color prints. Early experience with these prompted a switch to 5x7" format; this significantly reduced the number of flight lines and the number of photographs. Although the saving was more than offset by a doubling in the cost of prints (from 35 cents to 65 cents) and an increase in flight altitude from about 7500 feet AGL to about 10,000 feet, it was thought that the increase in ground coverage per photograph and the reduction in gross numbers of frames more than justified the expense.

The photos were acquired using three Nikon F3/T cameras and associated equipment. The cameras were fitted with motor drives and 250-exposure magazines. Use of long rolls of film minimized airborne cassette-changing time, but effectively restricted film choice to Kodak Vericolor III, a film with a nominal ASA speed of 160 normally used in studio and commercial work. Lenses with a focal length of 35mm were found to be most suitable; this length was short enough to permit photo acquisition at the proper negative scale (about 1:87, 000) without climbing to oxygen altitudes, and long enough to avoid the vignetting and airframe interference problems that tended to arise with wider-angle lenses. Ultraviolet filters were standard. Locally fabricated battery-powered intervalometers completed the camera suites, which were independent of aircraft power.

No camera mounts, in the usual sense, were employed. Two capscrews were fitted into holes drilled and tapped in the ends of the magazine; these together with an electrical connector on the motordrive unit formed a camera tripod with two adjustable legs. A bubble level was attached with carpet tape to the camera back. The camera rested on a sheet of quarter-inch-thick clear acrylic taped over the belly hole of the photo aircraft, with the lens protruding through a cutout. With the cutout centered in the belly hole, the photographer could be reasonably confident that no airframe interference would occur, and at the same time the camera could be freely rotated to correct for crab.

The photo format allowed a flight line spacing of 1.25 miles with approximately 25 per cent sidelap, and an interval of approximately .5 miles between photo stations to achieve 60 per cent endlap. Flight lines were laid out on 1:125,000 planimetric county road maps. These maps, produced by the Minnesota Department of Transportation, show the road and drainage scheme and the section line network--sufficient detail for orientation--and also bear latitude/longitude ticks at 5-minute intervals. The geographic coordinates were essential in flight planning, as LORAN radio navigation played an indispensable part in organizing the coverage. Standard meridians were established every 30 minutes of longitude, and the flight lines were laid out parallel to these at a spacing of 1.1 nautical miles--near enough to the desired 1.25 statute mile spacing.

Rental aircraft were used. Minimum requirements were (1) a belly hole for photography, and (2) LORAN capability. There were few enough aircraft meeting these specifications that scheduling was a continual preoccupation during the project. Most of the aircraft used were the common Cessna 172 type, renting for about $70 per hour with pilot, having an off-center belly hole about 7 inches wide and a foot long. Pilot training proved unproblematic; with a little practice--and some exceptions--the average commercial pilot proved capable of flying along LORAN lines within a tolerance of .1 nautical mile, and quickly learned the routine of changing LORAN settings while turning from one line to the next.

Inventory foresters received training in aerial photo procedures and took the bulk of project photography. Ground training covered project layout, recordkeeping, and camera manipulation; in-flight training accustomed photographers to handling the equipment, adjusting interval and crab, and keeping track of mission progress from a horizontal position in the cramped rear compartment of a light aircraft. A few photographers had airsickness problems, but for most the flight profile--chiefly cruising at around 11,000 feet MSL on clear days--was not troublesome. About four hours' ground training and two hours' flight training were enough to prepare a photographer to try his own hand. A checklist on the back of the photographer's clipboard reduced the chance of major omissions before takeoff. Photographers worked on every suitable photographic day from leafout to fall color change, weekends included.

Recordkeeping, in a project of this scope, was vital. Standardized flight logs were produced for photographers' use, and a double-labeling system was devised for film cassettes, one numbered label staying with the cassette from loading through final print inspection and the other being transferred to the flight log immediately following exposure. Flight logs were collated by county after each day's flight, and the lines flown each day were posted to a master map.

Processing and printing were performed under contract by a high-volume photo laboratory in the region. By providing this firm with a large quantity of photo work, we obtained the best of two worlds--consumer-level prices combined with an indispensable degree of custom attention to our needs. Standard consumer-grade mass-production printing without special attention is usually not good enough in aerial color negative photography: aerial scenes differ greatly from the general run of consumer photos processed on modern automatic machinery, and color balance usually varies to an unacceptable degree if no special filter setup is used.

Finished 5x7" glossy prints arrived in our offices in 250-exposure rolls together with the processed negative rolls from which they had been made. On receipt, the appropriate flight logs were consulted and a quick tracing of the actual flight line flown was mapped by inspection of the roll, to reveal any gaps in coverage. After this the individual frames were cut from the roll, arranged in flight lines, numbered, and filed. Negative rolls were also separated into flight lines, which were labeled and individually boxed for filing. Within any county we attempted to number all flight lines in order from west to east; within any flight line, photos were numbered from north to south. With the photos thus separated, finished index maps could be made showing actual flight line alignments and the center point of every tenth frame. Due to the urgency of the inventory project, much of the photography was used for its intended disturbance-checking purpose well in advance of final indexing.

ACCOMPLISHMENTS AND COSTS

By these means, over the 1986 and 1987 photo seasons and a portion of the 1988 season, the Resource Assessment Unit acquired some 49,000 photographs covering about 27,000 square miles in northern Minnesota, including all of Cook, Lake, St. Louis, Koochiching, Carlton, Itasca, Aitkin, Beltrami and Clearwater Counties, and the forested third of Becker County. Total cost of photographing these units was about $103,000, as itemized in Table 1. This amounts to a per-plot cost of $14 for the approximately 7500 commercial forest plots in the two units, and a cost of $4 per square mile for continuous stereo coverage.

Table 1. Costs of FIA photography project, 1986-1988.

Planning	$	3,500
Film	$	4,000
Aircraft/Pilot	$	44,500
Inflight hours	$	8,500
Processing & printing	$	38,500
Indexing & labeling	$	4,000
Total Cost	$	103,000

Interpretation of this photography resulted in avoided measurements on 996 plots in the Aspen-Birch Unit and on 1198 plots in the Northern Pine Unit, together covering the most heavily forested portions of Minnesota. At a $250 saving per avoided plot, this amounted to a net saving of more than $400,000, much of which was invested in new plot installation to strengthen the inventory.

Because this project represents the most recent full-coverage stereo photography in many areas of northern Minnesota, the products have been in high demand outside the Resource Assessment Unit. To date the Unit has filled orders totaling about 46,600 reprint frames for 66 agency and private customers, collecting $50,260 in receipts at prices we have set to cover costs of obtaining reprints and handling orders.

During the second half of the 1988 photo season the Unit flew still more 1:15,840 disturbance photography at the request of the North Central Forest Experiment Station, which is handling FIA inventory of the other two units in the state. Because of the more scattered distribution of the 1795 commercial forest plots in the Central Hardwood and Prairie Units, most of them were photographed individually, although one county received continuous coverage. Project cost for these two units was about $30,600, or just over $17 per plot. These photos were delivered to the North Central Station, where they are now in use.

LITERATURE CITED

Hackett, R. L., M.P. Meyer and N. L. Miller. 1983. 35mm aerial photography for analyzing forest land use change. In: Renewable Resource Inventories for Monitoring Changes and Trends: Proceedings of an International Conference. Aug. 15-19, 1983. Oregon State U., Corvallis. pp. 70-74.

Hahn, J. T., N.L. Miller and B. L. Essex. 1981. Application of growth simulation and small format photography to forest resource inventories. In: In-Place Resource Inventories: Proceedings of a National Workshop. Aug. 9-14, 1981. U. of Maine, Orono. pp. 947-950.

Meyer, M., S. Opseth, N. Moody and L. Bergstrom. 1982. Large-area forest land management applications of small scale CIR 35mm aerial photography. IAFHE RSL Research Report 83-2, College of Forestry, U. of Minnesota, St. Paul. 23p.

Meyer, M., N. Miller and R. Hackett. 1981. Application of 35mm aerial photography to forest land use change analysis in the upper peninsula of Michigan. IAFHE RSL Research Report 81-2, College of Forestry, U. of Minnesota, St. Paul. 19p.

A Study of Pinus Mugo-Stands in the Karwendel Mountain Region (Tyrol, Austria)

Dipl.-Ing. Josef Hackl
Dipl.-Ing. Erwin Knappitsch
Federal Environmental Agency, A-1090 Vienna

INTRODUCTION

A Colour-Infrared-Aerial-Photography assessment of the forest in the vicinity of Innsbruck (Tyrol, Austria) indicated that significant damage occurred locally in stands of *Pinus mugo*. The object of this study was to investigate the distribution, coverage and condition of *P. mugo* on the southern slope of the Karwendel mountain ridge.

MATERIAL AND METHODS

Large-scale (1:4000-1:5000) CIR-aerial photographs of the study-area were taken. An interpretation key was set up after thorough "in situ" inspection of individual patches of *P. mugo*.

Criteria chosen for an overall qualitative differentiation of *P. mugo* stands were as follows:

Condition 1:	(no or little discernible sign of damage): on average > 4 "needle years" found in the stand
Condition 2:	(frequent occurrence of damage symptoms): on average 2 to 3 "needle years"
Condition 3:	(heavily damaged or dead): on average < 1 "needle year"
Remnant stands:	Patches or individuals of *P. mugo* remaining from previous stands.

Qualitative and quantitative assessment of the status of *P. mugo*-stands was carried out with an analytical stereometric interpretation device. Data were incorporated into a 3-dimensional graphical model of the study area.

RESULTS

The investigated area amounted to 890 ha, with *P. mugo* covering 782 ha.

The relative frequency of stand area assigned to condition 1, 2 and 3 amounted to 63.8%, 33.6% and 2.6%, respectively (fig.1). Heavily damaged stands are typically found above 1,600 m altitude.

The occurence of damage symptoms increases with altitude (fig.2).

Remnants of previous stands cover a total of 64 ha and are largely restricted to altitudes of more than 1,900 m above sea level.

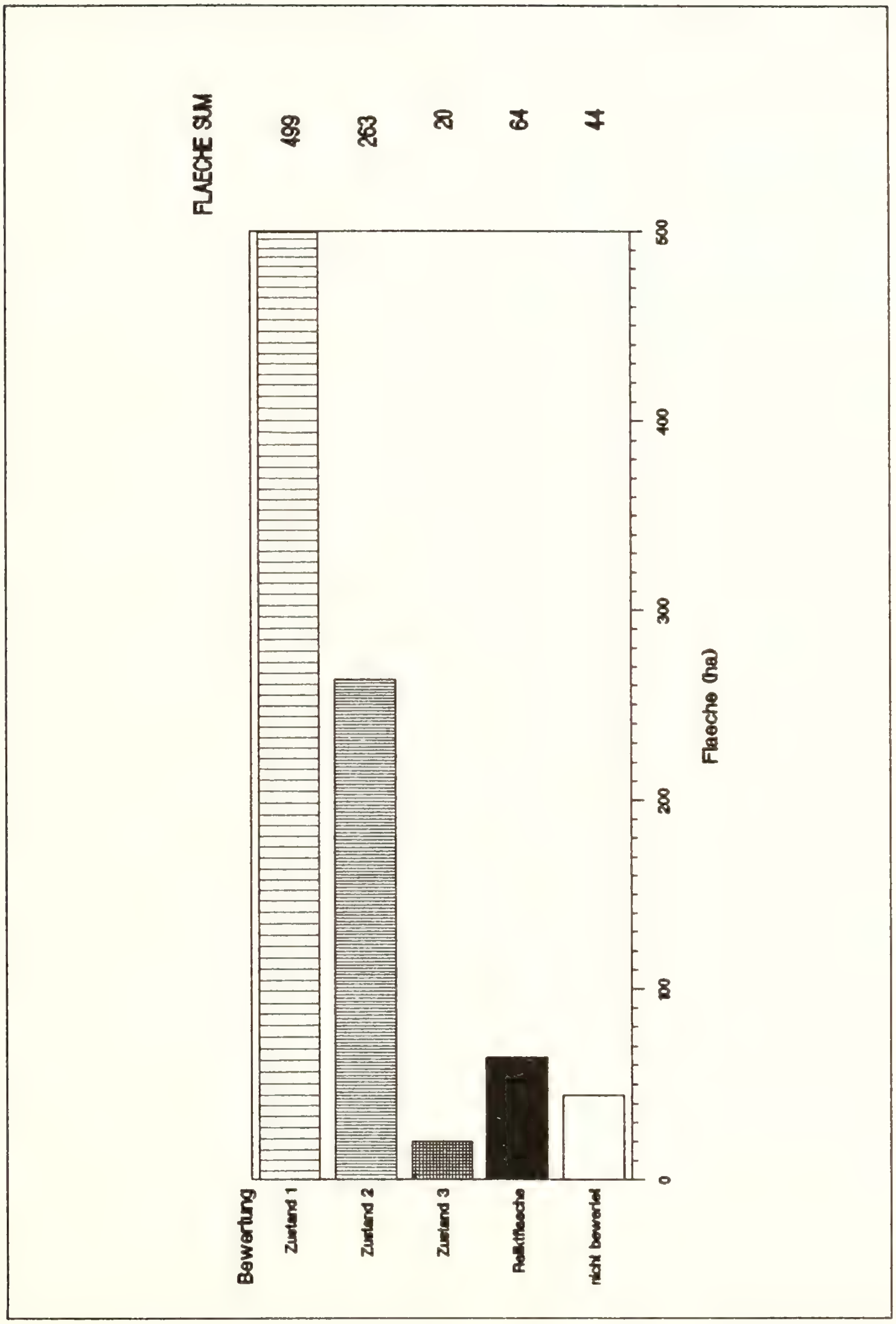


Figure 1. Vertical distribution of *P. mugo* stands with respect to symptoms of damage.

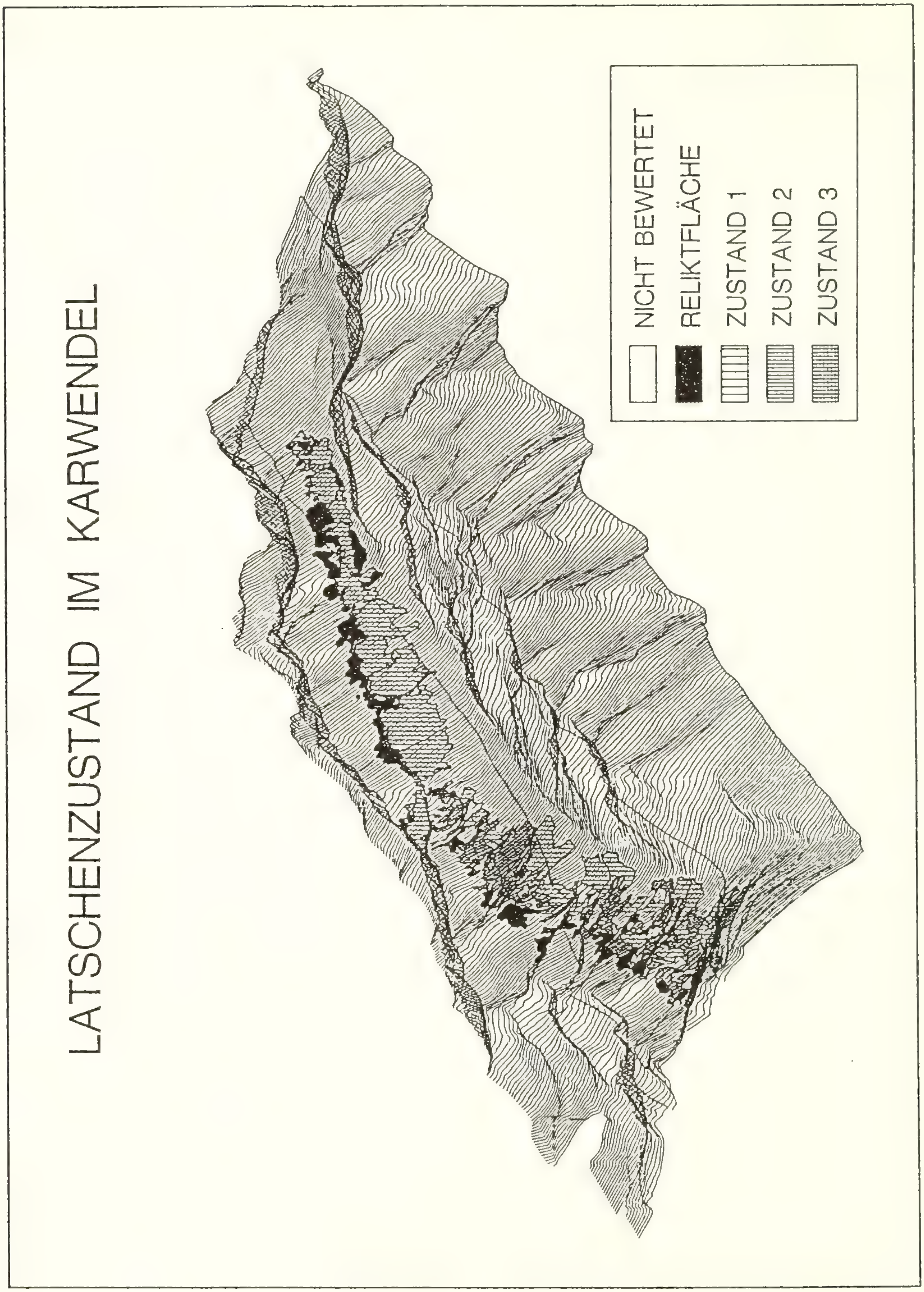


Figure 2. Areal distribution of *P. mugo* stands with respect to symptoms of damage.

DISCUSSION

Reasons for the apparent decline of P. *mugo*-stands in high altitudes are manifold. They probably include the climatic changes in the mid 19th century (Kral, 1988), as well as the current impact of atmospheric pollutants, particularly ozone (Gruber, 1988).

To document whether and to what extent the upper limit of *P. mugo* has significantly changed during the last decades, areal photography surveys conducted in the fifties are presently being evaluated.

Concurrently the impact of mechanical damage caused by avalanches and rock-fall in steep dry debris slopes on the areal distribution and "vitality" of *P. mugo*-stands is investigated.

REFERENCES

Gruber F., 1988: Waldbauliche Beurteilung von Schäden in Latschenbuschwäldern (Karwendel). Diplomarbeit an der Universität für Bodenkultur, Institut für Walbau.

Kral F., 1988: Pollenanalytische Untersuchungen zur Entwicklungsdynamik der Latschenbestände im Karwendeltal (Tirol), Cbl.f.d.ges.Forstw. 105, pp. 23-35.

Österreichisches Bundesinstitut für Gesundheitswesen, Hrsg.: Stadtvegetation Innsbruck 1977 - 1987. Im Auftrag des Stadtmagistrats Innsbruck. Fric - Verlagsgemeinschaft - Manz.

Forest Health and Decline
Moderator: T. Cunia

Joseph E. Barnard, Environmental Health Concerns: A Role for Forest Inventory and Monitoring

John P. Jastrzembski and James P. Barrett, Monitoring the Health of White Pine in New Hampshire (1959-1988)

Michael Koehl, National Inventories and Inventories of Endangered Forests in Europe

Paul C. Van Deusen and Joseph E. Barnard, Indicator Plants in Forest Health Surveys

ENVIRONMENTAL HEALTH CONCERNS: A ROLE FOR FOREST INVENTORY AND MONITORING

Joseph E. Barnard, Program Manager
National Vegetation Survey/Forest Response Program
Southeastern Forest Experiment Station, P.O. Box 12254
Research Triangle Park, North Carolina 27709

ABSTRACT

Factual information and correct interpretation of forest damage in the context of atmospheric pollution and other environmental issues is critical to public policy, land management and individual ownership decisions. The USDA Forest Service is developing a forest health monitoring program. The first step was the implementation of a forest health survey, called the Visual Damage Survey. Data collection activities began during the summer of 1988 on 400 plots located in three separate ecosystems. Full implementation of comprehensive forest health monitoring is the next step. This is envisioned to be a multi-tiered process to (1) detect unexpected deviation from established baseline conditions or trends, (2) identify cause, and (3) define basic relationships sufficient to predict consequences. Consisting of three increasingly detailed levels: detection monitoring, evaluation monitoring, and research monitoring, this proposed monitoring program will be an integral part of the Forest Service's Forest/Atmosphere Interaction Program proposed under the Forest Ecosystems and Atmospheric Pollution Research Act of 1988 (PL 100-521).

INTRODUCTION

Americans are quick to react to any issue or problem that appears to threaten the forest. They are concerned that the nation maintain a healthy, vigorous forest resource that will satisfy their aesthetic, recreational, and consumptive appetites. The injury and death of trees in the forest by a variety of vectors is part of the natural rhythm of the ecosystem. When injury or death occurs as an isolated incident, the natural resiliency of the system enables the forest to maintain an overall healthy state. However, when large numbers of trees or extensive areas of the forest are impacted, there is reason for concern.

During the 1980's, foresters have been confronted with the issue of "acid rain" and the questions of damage to forests by a variety of factors degrading the environment. Existing inventory data and measurement procedures have served a broad and useful monitoring function by alerting us to changes in growth trends and other measures of forest condition. However, we are unable to fully address the concerns and provide more specific responses because these inventories were not designed with monitoring and hypothesis testing in mind. Such forest resource information was not previously gathered on the regional scale.

None-the-less, factual information and correct interpretation of forest damage in the context of atmospheric pollution and other environmental issues is critical to public policy, land management and individual ownership decisions. Since these decisions have important social and economic impacts, it is essential for the decision makers to have good information on what is happening to the forest with respect to specific problems or issues. Furthermore, estimation of the future condition of the forest under a variety of alternative scenarios provides focus to the decision-making process.

FOREST HEALTH INVENTORY AND MONITORING

How does one acquire the necessary information? Enter: inventory and monitoring.

Let me begin by defining the two concepts--inventory and monitoring. Inventory is the enumeration of pertinent data such as volume, increment, mortality, etc., for the assessment of particular characteristics such as productivity, health, or damage at a particular point in time. Repeated continuous inventory implies the regular enumeration of particular characteristics in such a way as to provide a check on the changes taking place in that characteristic. Monitoring is the repeated recording or sampling of pertinent data for the comparison of that data to a reference system or identified baseline. Monitoring always involves the determination of changes over time and usually also involves interpretation with respect to the reference or baseline. The principal differences are in:

1. The degree and intensity of sampling and parametric detail,
2. The frequency of sample collection and
3. The sophistication of instrumentation used.

A rather obvious point to note is that renewable resource populations typically consist of large numbers of individual elements distributed over appreciable areas in such a manner that it is either logistically impossible or economically impractical to conduct a complete census. Sampling technology is an indispensable part of the process. Forest resource inventory/monitoring in the United States is based on statistical sampling methods. The few exceptions involving extremely small populations are of no interest in the present context.

Both inventory and monitoring activities produce information describing the current or past resource situation. Decision makers often need an indication of the future situation in order to develop or evaluate alternative courses of action. This necessitates the ability to project information forward in time or space. The human mind is remarkably capable of envisioning possible futures. However, when dealing with matters of great importance involving many individuals in the decision process, procedures that can be observed, used, and evaluated by others are necessary. The most common means of achieving this is by computer modeling.

The current assessments of atmospheric deposition impacts on forest health in the United States rely heavily on modeling. The lack of specific symptoms has forced scientists to evaluate existing inventory data collected for other purposes and develop assumptions for testing with models developed for other purposes. A major effort of the researchers in this field is the development and use of appropriate models and complete forest ecosystem simulators.

OVERVIEW OF SOME CURRENT PROGRAMS

Canada

Forestry Canada (formerly Canadian Forestry Service) implemented an Acid Rain National Early Warning System (ARNEWS) in 1984. The objectives of that program are to: (1) detect the possible damage to forest trees and soils caused by air pollution or identify the damage sustained by Canadian forests that are not attributable to natural causes or management practices, and (2) conduct long-term monitoring of vegetation and soils to detect future changes attributable to air pollutants in representative forest ecosystems.

To date over 110 permanent plots have been established throughout forest areas of Canada. These plots will be maintained for at least 20 to 25 years to monitor: (1) condition and changes in condition of the forest stand, (2) presence and fluctuation of biotic and biotic factors that affect the condition of the forest (insects, diseases, stand changes, temperature, etc.), (3) changes and symptoms that indicate factors not identified in item (2) that could be early signs of air pollution damage, and (4) effects of air pollution on the condition of the various economically important tree species. The Canadians have not reported on the early findings of this monitoring effort; but the data and field observations related to this program are used repeatedly in the on-going Forest Pest Management Protection program of the federal and provincial governments.

Europe

The United Nations Environment Program/Economic Commission for Europe began a unified program of "forest damage" monitoring throughout Europe in 1986 under the auspices of the International Convention on Transboundary Air Pollution. The objectives are to: (1) establish a harmonized system for assessment and monitoring effects of air pollution on forest health, and (2) operate a system for monitoring forest damage from air pollution. The program recognizes the need for several levels of activity: (1) large-scale representative surveys, (2) intensive studies on permanent plots, and (3) special forest ecosystem analyses. To date, most attention has been given to implementing the first activity, representative surveys, on a Europe-wide scale.

The representative surveys are conducted by each participating country according to certain minimum standards. Sample points are laid out on 4 km by 4 km or 16 km by 16 km grids, depending on the size of the country. This system has been very successful in tracking the year-to-year variation and progression of very generalized damage symptoms throughout Europe.

These data along with some analysis are reported annually. Those responsible for the data collection in each country meet each year to review the results and develop appropriate recommendations and conclusions. It is now widely accepted that these symptoms alone do not enable conclusions about the cause of the damage. To attribute causality, the other two levels of activity--intensive permanent plot studies and ecosystem studies--are necessary. Efforts at implementing the other levels of the program are being considered.

Canada and the United States participated in the development of this European monitoring program. The Canadian ARNEWS Program employs similar crown condition rating procedures, and regional United States surveys of forest health have used the European crown condition classifications.

USDA Forest Service

Currently, the USDA Forest Service gathers data about the forest resource and its health in various ways. The two activities most germane to this discussion are the periodic Forest Inventory and Analysis (FIA) surveys and specific insect and disease surveys conducted or coordinated by the State and Private Forestry division of Forest Pest Management (FPM). These two activities supplement each other. The periodic forest resource inventory records tree mortality as well as many other attributes on a network of permanent ground plots while the specific insect and disease surveys, based on aerial and/or ground samples, record short-term changes in pest activity.

The Forest Service has been developing a forest health monitoring program during the past five years. A major impetus for this development has been the concern about documented and potential effects of air pollutants on forests in North America and Europe. Concern over the condition of forests in response to traditional stresses (e.g., insects and diseases) and multiple, interacting stresses has also been an important factor in shaping the program.

In 1984, we implemented a National Vegetation Survey to produce information on the state of United States forests with respect to air pollution impacts. The National Vegetation Survey became one of six research cooperatives in the jointly sponsored Forest Response Program in the National Atmospheric Precipitation Assessment Program (NAPAP). The sponsors are the U.S. Environmental Protection Agency (EPA), the USDA Forest Service and the National Council of the Paper Industry for Air and Stream Improvement.

Several widely attended planning meetings involving Forest Service, university, and industry scientists resulted in the conclusion that it was premature to implement a national survey at that time. We simply did not know enough about low level pollution effects and symptoms, and we did not have sufficient funds to undertake a survey that was not highly focused or limited geographically. Instead, we began to monitor forest condition in limited forest areas, to conduct surveys across pollution gradients to improve our ability to interpret monitoring results, and to study monitoring techniques and analysis methods.

Since 1984, the Forest Response Program has sponsored a number of monitoring-related research activities. These include more detailed analyses of existing FIA plot remeasurement data, analysis of increment cores collected from trees on FIA plots; the establishment of plots along elevation gradients in the high elevation Spruce/Fir forests of the Appalachian Mountains; the selection and measurement of FIA plots along two deposition gradients in the North; the selection and measurement of a series of Sugar Maple sites in the eastern Canadian provinces and the eight US states that border these; the continuous annual measurement (since 1984) of a set of spruce plots in Massachusetts, New York, New Hampshire, and Vermont; the identification, editing and analyses of existing southern pine research plot data sets from all southern states.

Analyses of existing data coupled with scientist interactions with inventory/ monitoring groups in Canada and Europe has resulted in the development and initial implementation of a forest health survey, called the Visual Damage Survey. The available plot systems of the Forest Service's FIA and FPM groups provide the spatial capability to collect the data. Statistical details to enable maximum sampling efficiencies are being evaluated during the initial phase of implementation. Plot data collection procedures were developed by a committee of scientists and practitioners representing Canadian and European interests plus state, university and Forest Service groups. A field manual developed from these inputs identifies tree and plot data gathering procedures that are consistent with the UN-ECE Europe-wide International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests and with Forestry Canada pest management studies including ARNEWS. The current survey does not include several foliar chemical procedures used in Canada. Instead the forest health survey places emphasis on hands-on examination of foliage from within the tree crown in addition to ocular classification of visible conditions. This survey also attempts to assess the condition of the tree root systems using a unique sampling system developed in the South. The inclusion of these two additional procedures was strongly supported by Canadian and European reviews.

Data collection activities began during the summer of 1988 on 400 plots located in three separate ecosystems: Spruce/Fir in the North, hardwoods in Pennsylvania, and natural Loblolly pine in the southern Piedmont. These ecosystems were selected because earlier research had indicated their significance and because they provide wide geographic coverage. Current plans are to expand this network in a logical manner as inputs from other research and survey efforts indicate. During the 1989 field season crews will revisit a portion of the sites surveyed in 1988 and add plots in plantations of Norway spruce in the North and in natural Loblolly pine stands in the coastal plain of North Carolina. We are evaluating ways to accelerate the implementation of this survey using crews provided by cooperators through a partnership effort.

Initial evaluations of the data from 1988 show the visible condition of most of the tree crowns evaluated to be good. The presence of potentially growth inhibiting fungi in the root samples of over 25% of the loblolly pine sampled was an unexpected finding which has generated additional research interest. The levels of root pathogen activity in the hardwoods and spruce trees sampled were very low (less than 5% incidence).

A separate but related survey is being jointly conducted by the Forest Service, the states, Forestry Canada, and the provinces. This is known as the North American Sugar Maple Decline survey. It consists of 166 sites in the sugar maple region. At each site observations of tree condition along with mensurational data on the trees is being collected. The first measurements were made in 1988 and continued efforts have been agreed upon by all cooperators for 1989 and 1990. This work relates quite nicely to the forest health survey network plot procedures; but its sites are not associated with either the FIA plot network or the Canadian ARNEWS plot system. This sugar maple survey represents the type of detailed response required when a specific problem, i.e. maple decline especially in sugar bushes, is being investigated. Much of the information developed will be available for integration in the analysis of forest health. The cooperation accomplished in this effort is a model for future partnerships.

Future Forest Service Plans for Forest Monitoring. Full implementation of comprehensive forest health monitoring is the next step. This will build on past research and experience gained from pilot tests, specialized forest health monitoring, operational forest inventories, and forest pest evaluation activities.

The current Forest Inventory and Analysis and Forest Pest Management inventories of the USDA Forest Service will provide the base monitoring network. This broad, statistically based network will be supplemented by several additional components including forest pest surveys, targeted monitoring of priority areas, additional Forest

Inventory and Analysis plot sampling, and information from intensive research sites. The key concept is that each component must be integrated into the base monitoring network; this will mean that data from each component will have the spatial and temporal attributes associated with national, regional, and state forest statistics.

Comprehensive forest health monitoring will be done within existing Forest Service responsibilities: the FIA inventory, National Forest System's inventory, FPM surveys, evaluations conducted by the Forest Service and states, and Research investigations on intensive monitoring sites. Recent pilot project implementation of cooperative forest health monitoring of spruce-fir and sugar maple areas began the movement toward such a monitoring system.

Forest health monitoring will be a multi-tiered, long-term process to (1) detect unexpected deviation from established baseline conditions or trends, (2) identify cause, and (3) define basic relationships sufficient to predict consequences. Each successive tier requires progressively more detailed and costly information. The three increasingly detailed monitoring levels are detection monitoring, evaluation monitoring, and research monitoring. All are essential to a successful monitoring system. The three levels are described below:

Detection monitoring (or "routine" monitoring) is the first and most extensive level of this system. This level will represent a geographically-based network of permanent plots coupled with routine pest surveys distributed throughout the forests of the United States. The spatial basis for these permanent plots will be the forest inventory (FIA) permanent plot network maintained by the USDA Forest Service. The FIA network would be augmented with additional plots to represent all forest lands. From this augmented network, a subset of "sentinel plots" will be selected and routinely visited at intervals ranging to several times per year to biennially. The amount of information collected will be greater than on the normal forest inventory plots. Sampling will be conducted by Forest Service crews and by crews provided by cooperative partners.

Information from the sentinel plots will be supplemented by information collected during routine forest pest surveys (by Forest Service FPM and by state agencies) and other specifically focused monitoring activities. This information will be spatially linked (through use of the Geographic Information System) to the network of sentinel plots, thus providing a more complete annual estimate of forest condition. The objectives of this level of monitoring are to:

1. Enable systematic observation of key monitoring elements that "describe" forest condition and assist in the early detection of change "signals" calling for more detailed evaluation.
2. Provide data for annual forest condition (health) statistics.
3. Activate evaluation monitoring (level two) in a planned and coordinated manner.

Data obtained from these plots and other sources of detection-level information will be managed and evaluated by a staff of experts within the Forest Service. Monitoring results may be reported in several ways including: time series of the basic observation variables; synthetic "indices" of condition, and appropriate summaries for use by the states, other federal agencies, cooperators and the research community. This level of monitoring will be the primary linkage to EPA's proposed Environmental Monitoring and Assessment Program (EMAP).

Evaluation monitoring (or "ad hoc" monitoring), the second level, is activated by detection monitoring results. When detection results identify areas or problems of concern, a multidisciplinary team will determine specific evaluation needs and undertake activities such as additional surveys, site or area specific evaluations, and more detailed monitoring. The objectives of these detection-activated evaluations are to:

1. Clarify concerns raised by detection monitoring, delineate problem extent, identify cause (if possible), and make control recommendations as appropriate.
2. Identify research hypotheses if cause can not be identified.
3. Provide the additional data needed to link evaluation of these relatively broad problems with the detailed, site-specific research monitoring (third level) undertaken to define and predict causal relationships and effects.

Examples of the more detailed monitoring and evaluation at this level are the current spruce-fir and sugar maple surveys.

Research monitoring (or "investigative" monitoring) is the third level of this system. It provides very detailed information on all components of the forest ecosystem and supports research on specifically identified intensive research sites where both special and continuing long-term studies are conducted. These sites will represent key forest ecosystems throughout the United States. The purpose of these detailed monitoring sites is to supplement detection and evaluation monitoring and to support mechanistic research to identify, describe, or model tree, stand and forest processes in ways that (1) increase basic understanding of causal relationships and (2) enable explanations or projections of observations in the other levels of this monitoring system. Monitoring at this level provides data for ecosystem research undertaken to better understand and predict causal relationships and rates of changes in forest condition. Research natural areas, experimental forests, intensive research sites such as those of the Forest Response Program, and locations such as Coweeta and Hubbard Brook are examples of this level. Forest Service Research Stations and the universities will be responsible for these sites. Studies will be needed to develop appropriate linkages between these sites and the more extensive aspects of the detection and description data and analyses. Modeling activities would help provide this linkage.

The scope of forest health monitoring will be such that it addresses forest health relative to effects of both naturally occurring factors such as fire, forest pests, forest succession, and drought, and unnatural biotic and abiotic (anthropogenic) factors such as introduced pests, air pollution, acidic deposition, and global warming. The term "forest health" describes forest ecosystem resilience and productivity relative to public values, needs, and expectations. A healthy forest can be described by different standards, each related to differing management objectives for particular forested areas. The first step toward evaluating forest health is describing forest condition. The three levels of monitoring described above will provide a thorough description of forest condition as well as sufficient information on causality to enable evaluation of forest condition in terms of forest health.

An index is a composite mathematical function that summarizes the appropriate information as an abstract term. A stock market index which retains the same abstract meaning in different economic markets, even though based on different measurements and different functional expressions, is a helpful example. Some have suggested that a forest health index could be constructed as a function of indicators, a classification scheme and other environmental variables. The concept needs considerable attention and developmental thought before its application can be recommended. The identification of indicators is underway. Considerable work remains to resolve a number of specific questions regarding what to measure, measurement standards, data analysis procedures, etc. However, there is general consensus that detection monitoring indicators should eventually include measures of forest growth, forest mortality, forest distribution and structure, canopy condition, soil condition, and surface water condition. Debate concerning the use of these and other appropriate information in one or several "indices" of forest health needs to begin.

Relationship to Other Proposed Forest Service Research Activities. This proposed monitoring program will be an integral part of the Forest Service's Forest/Atmosphere Interaction Program proposed under the Forest Ecosystems and Atmospheric Pollution Research Act of 1988 (PL 100-521). The goals of the Forest/Atmosphere Interaction Program are to (1) determine which processes in forest ecosystems are sensitive to changes in the atmosphere, (2) determine how future atmospheric changes will influence the structure, function, and productivity of forest ecosystems, and (3) to determine the implications for forest management to sustain forest health and productivity in a changing atmospheric environment. The proposed program under PL 100-521 concentrates on the relationships between forests and the chemical environment. However, we recognize that much of the emphasis must be on ecosystem responses to multiple stresses, and it will often not be possible (or desirable) to segregate research on the various influences of the physical, chemical, and biotic environments.

SUMMARY

Both the public and forest-land management agencies are concerned about the health of the forest resource. The first step in the management decision process is the gathering of useful information. In this paper, I have noted the distinction between inventory and monitoring and discussed the general approach to their implementation for forest damage assessments. Inventory and monitoring need to be conducted regularly to provide such information. The uses of the information developed will be as varied as the data itself. Now we have an opportunity to design and implement a monitoring program as a legacy of research on the effects of acidic deposition on forests. This will require the best efforts on many of us in the field of forest resource inventory.

ACKNOWLEDGMENTS

The author gratefully acknowledges the important contributions of David L. Radloff and Robert C. Loomis to the section on Forest Service activities.

Monitoring the Health of White Pine In New Hampshire (1959-1988)

John P. Jastrzembski
Chesapeake Bay Watershed Forester
Maryland Dept. of Natural Resources
Forest, Park and Wildlife Service

and

James P. Barrett[1]
Professor of Forest Resources
University of New Hampshire
Durham, New Hampshire 03824

ABSTRACT

Survival rates were determined using the logistic function and data collected from 45 permanent even-aged, unthinned, eastern white pine (Pinus strobus L.) growth plots in New Hampshire. Three-year survival rates were predicted based on explanatory variables of the 3-year tree basal area average annual growth and the ratio of dbh/msd.

Environmental and biological factors influencing tree growth and survival are summarized and discussed. Increment cores indicate that 18 percent of the trees sampled exhibit symptoms of root and butt wood decay. 38% of the foliage sampled exhibited symptoms of ozone damage. Less than 2% of the trees showed evidence of attack by the red turpentine beetle (Dendroctonus valens LeC.) or signs and symptoms of white pine blister rust (Cronartium ribicola Fisch. Rabenh).

General growth relationships are tabulated and discussed. Ozone damage affecting greater than 20% of the foliage and wind associated crown abrasion appeared to decrease tree diameter growth.

---

[1] Former Research Assistant, University of New Hampshire, currently with the Maryland Department of Natural Resources, Forest, Park and Wildlife Service, and Professor of Forest Resources University of New Hampshire respectively. Scientific Contribution No. 1630 from the New Hampshire Agricultural Experiment Station.

INTRODUCTION

White pine stands, just as any forest community, are dynamic ecosystems. Trees are constantly competing for survival with adjacent trees for space, nutrients, moisture, and sunlight. Competition occurs among the influences of the physical environment, insects and diseases. This survival process is repeated continuously until each tree has succumbed to competition, environmental factors, age or some combination of these influences (Buchman et al., 1983).

There is limited information on tree health and survival for eastern white pine in New England (Jastrzembski, 1989). With health and survival information, managers can more accurately predict forest growth and yield. Survival models are often used to help schedule stand management activities. Stands of trees with high mortality rates predicted should be considered high risk stands, perhaps because of susceptibility to insects and disease outbreaks. These stands should be salvaged or otherwise treated to improve health and lesser risk before stands composed of trees with low predicted mortality rates (Hamilton and Edwards 1976).

Tree health models also provide valuable information pertaining to specific relationships which influence tree survival such as stand density, site index, tree age, and crown position. Likewise, these relationships are valuable in forest management planning and in developing strategies to control mortality (Hamilton and Edwards, 1976). The objectives of this paper are to:

(1) Present one of a series of statistical equations developed, based on the logistic function, to predict survival in naturally regenerated, even aged, white pine stands.
(2) Summarize some of the more important biotic and abiotic factors which influence white pine vigor, health, and survivability.

Field Data

In order to monitor tree health and survival, permanent plots are generally needed (Clutter et al., 1983). For this study the University of New Hampshire growth plots were used. These 45 plots, monitored since 1959, consist primarily of naturally regenerated, even-aged, originally unthinned, white pine trees. The plots range in size from 1/20 acre to 1/5 acre. The plots are considered typical of white pine stands in New England: the average site index is 67 with a range of 56 to 82 based on breast height age; basal area in 1988 averages 149 square feet per acre with a range of 46 to 266; the average stand age at breast height in 1988 was approximately 78 years ranging from 52 to 130 years. In the past decade some of the stands have been thinned.

Tree Mortality

Most of the plots have been measured at three year intervals from 1959 to the present. Therefore, the logistic model was used to estimate survival over a three year period during a time when the stands were unthinned. Two consecutive 3-year growth intervals (three sets of measurements) are needed to predict survival. The first 3-year interval is used to measure the current tree growth rate. The second set of measurements is used to gather stand and tree information. Survival or mortality is recorded at the end of the second growth interval.

Using logit analysis, a large number of situations can be modeled. The variables can be nominal (indicator or dummy), or interval (continuous or non-continuous). The variables can also be expressed as transformations and interactive terms (Kleinbaum et al., 1982). Maximum likelihood estimation assures the large sample properties of consistency and asymptotic normality of the parameter vector so that conventional statistical tests of significance are applicable (Capps 1983). Explanatory variables which were used for predicting mortality include:

crown class
age
site index
stand basal area
drainage class
soil texture class
dbh / msd
dbh
mean stand diameter
mean tree basal area annual growth (bag)

The dependent variable (3-year survival) is binary taking on a value of 1 for trees which live and 0 for trees that die. The analysis was based on 4036 observations, including several observations on the same tree.

Several models were developed using the logistic equation. The model discussed here represents one of the more practical ones used to estimate 3-year survival rates. This model is based on independent variables of individual tree 3-year basal area average annual growth in square feet (bag), and the ratio of tree dbh divided by the plot mean stand diameter (dbh2msd2). Table 1 includes the number of observations, chi^2 test, variable coefficients, standard errors, t-values, level of significance, and the means variable values for this model.

Table 1. Logistic regression estimates for 3-year survival. Three-year survival is expressed in terms of the Z-index. Explanatory variables are the most recent 3-year basal area annual growth (bag) expressed in square feet, and the growth ratio of dbh/msd (dbh2msd2).

```
    Logit Estimates                                    Number of obs   =     4036
                                                       chi2(2)         =   455.14
    Log Likelihood = -728.83084                        Prob > chi2     =   0.0000

    Variable    Coefficient   Std. Error      t        Prob > it         Mean
--------------------------------------------------------------------------------
     survive                                                          .9363231
--------------------------------------------------------------------------------
         bag     88.05394    17.181751      5.125        0.000        .0114444
    dbh2msd2     4.368086     .3759883     11.618        0.000        1.009037
       _cons    -1.320361     .2557985     -5.162        0.000               1
--------------------------------------------------------------------------------
```

Note that the chi^2 test indicates significance (<.01%). The variables have large t-values, indicating a significant relationship with three-year survival. McFadden's alternate correlation coefficient[2] is 0.31 indicating a strong relationship between the explanatory and dependent variables.

The resulting equation is:

$$Z = -1.320 + 88.0539\,(\text{bag}) + 4.3681\,(\text{dbh/msd})$$

Then the probability of tree mortality is predicted by the equation:

$$p = \frac{1}{1 + e^{z}}$$

The signs of the coefficients are positive, indicating that as both the tree basal area annual growth and the ratio dbh/msd increase, the survivability of the individual tree will increase. The mean dbh/msd for live trees is 1.032 and 0.663 for dead trees; the mean tree average annual basal area growth is 0.012 square feet for live trees and 0.001 for dead trees.

Graphical representation of predicted survival versus dbh/msd (Figure 1a) indicates that the individual trees survival rate approach one for trees which are larger than the mean stand diameter. Trees which are smaller than the mean dbh have lower survival rates. However, the smallest trees still have almost a 50 percent chance of surviving the next three years, indicating that the chances for a tree to die over the course of one three year period are rather small. Figure 1b represents the predicted survival versus 3-year tree basal area annual growth. Survival rates are greatest for trees with the greatest growth rates and poorest for trees with little growth. A tree with a high growth rate is almost certain to survive unless struck by lightning or attacked by a bulldozer. The response for a tree with a low growth rate, however, is highly variable with survival ranging from about .5 to approaching 1.0.

---

[2] When using McFadden's R-square, a value of 0.33 indicates a relatively good relationship between the explanatory variables and the dependent variables

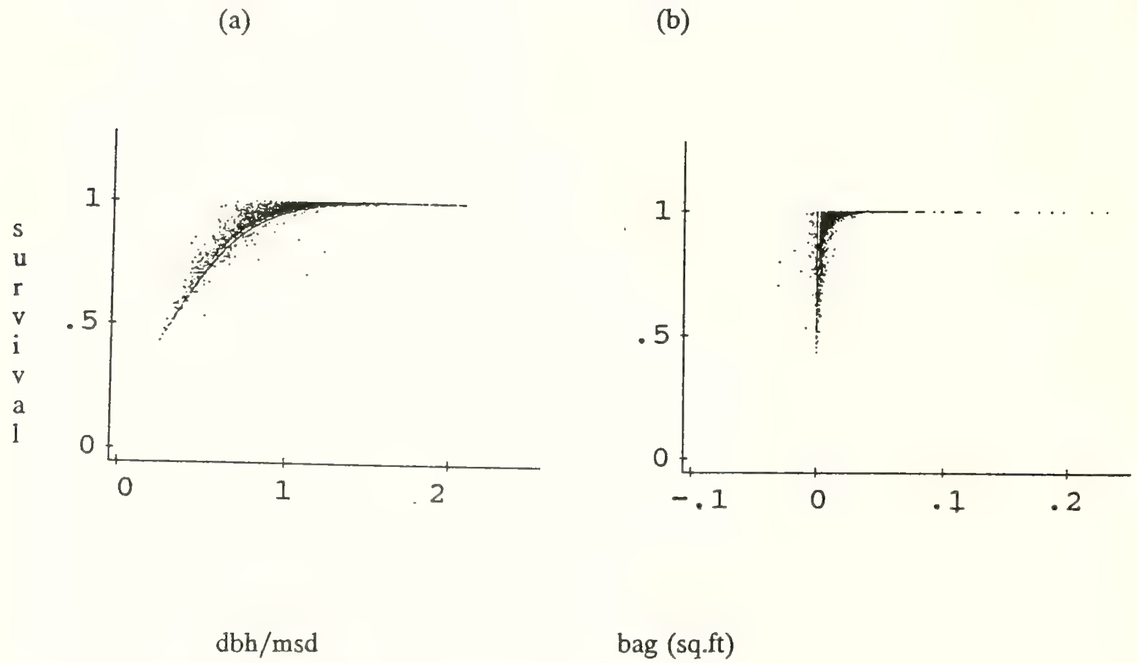


Figure 1. Predicted 3- year survival rates versus (a) dbh/msd and (b) 3-year tree basal area average annual growth (bag) in square feet.

Environmental and Biological Factors Influencing Health

The second part of this project, involved summarizing some of the more important biotic and abiotic factors which influence white pine health, vigor, and survival. Factors included white pine blister rust (C. ribicola (Fisch.) Rabenh), red turpentine beetle (Dendoctonus valens LeC), root and stem decay pathogens, ozone pollution, and crown abrasion.

Results of the 1988 field survey indicate that less than 2 percent of the trees, based on all the living trees on the 45 plots (n = 757), exhibited signs or symptoms of blister rust or beetle activity.

One hundred and sixty one trees were cored to determine the presence or absence of decay fungi. Thirty-five samples showed evidence of decay or discoloration. These samples were placed into cultures to isolate the fungi. Unfortunately, results were not conclusive due to contamination or sterility of the samples. However, results indicate that 82.6 percent of the trees were apparently healthy and that 17.4 percent exhibited signs of decay. Of the trees sampled, 9.3 percent, 6.2 percent, and 3.7 percent exhibited signs of brown rot, white rot, and red stain, respectively. However, failure to isolate the pathogen(s) precludes identification of the causal agent. Statistical tests determined that there were no significant differences between presence of the fungi and site index, stand age, or basal area.

Ozone pollution damage was detected on 38 percent of the foliage collected. Sources of emitted air pollution, weather patterns, and topography most likely account for the location of ozone pollution damage.

Trees were placed into two categories of ozone pollution damage levels (less than 20 percent and greater than 20 percent). It was determined that trees in the less than 20 percent category grew in average annual diameter growth (AADG) 8 percent faster than trees in the greater than 20 percent category (Table 2).

Table 2. Most recent 3-year average annual diameter growth (AADG) according to plot levels of ozone pollution damage (<20%, or > = 20%) for the University of New Hampshire white pine growth plots.

	PLOT OZONE DAMAGE LEVELS[a]	
	< 20 %	> 20 %
AADG[b]	.100	.093
standard deviation	.087	.086
observations[c]	110	672

[a] Percentage of foliage sampled per plot which exhibit symptoms of ozone pollution damage.
[b] Most recent three-year average annual diameter growth measured to the nearest tenth inch.
[c] The total number of observations in each category.

When average annual diameter growth (AADG) was compared to crown friction and crown class, the results were as anticipated (Table 3). As crown position increases, the AADG increases; as crown friction increases, the AADG decreases. Thus, dominant trees exhibiting no signs of crown friction grow faster than all other matrix combinations. Overtopped trees will most likely be suffering from moderate to high levels of friction.

Table 3. Summary of the average annual diameter growth (AADG) for the University of New Hampshire-white pine growth plots categorized according to crown abrasion and crown class for the most recent 3-year growth period.

		CROWN FRICTION			
CROWN CLASS		NONE[a]	LOW[b]	MODERATE[c]	HIGH[d]
	AADG[e]		.100	.017	.033
OVERTOPPED	sd.[f]	-----	.094	.023	.036
	obs.[g]		2	2	13
	AADG	.083	.025	.027	.027
INTERMED.	sd.	.042	.035	.031	.035
	obs.	4	8	15	46
	AADG	.116	.097	.066	.049
CO-DOM	sd.	.084	.093	.067	.045
	obs.	32	86	99	79
	AADG	.136	.128	.103	.063
DOMINANT	sd.	.095	.128	.103	.063
	obs.	139	141	68	19

[a] Trees did not exhibit any signs of crown abrasion.

[b] One of the four sides of the crown showed signs of twig and branch breakage due to crown friction.

[c] Approximately two sides of the live crown exhibited signs of crown abrasion.

[d] Approximately three or more sides of the crown exhibited signs of crown abrasion.

[e] Average annual diameter growth measured over the past 3-years. Measurements are recorded to the nearest 0.1 inch.

[f] The standard deviation.

[g] Based on the remaining living trees.

CONCLUSION

In conclusion, eastern white pine individual tree mortality and survival rates can be estimated using logistic model. The model described relates 3-year survival rates to dbh/msd and 3-year tree basal area average annual growth. Generally, mortality rates were low. Trees with high growth rates have a very high probability of surviving. The response is highly variable for slow growers: some die and some survive for a long time. Perhaps their ability to survive depends on root grafting with other trees.

Environmental and biological factors were summarized and discussed. Some biological variables, however, such as genetic variation in resistance to ozone damage, were not measured.

General growth trends were tabulated and discussed. It was determined that ozone levels (greater than or equal to 20%), crown position, and crown friction influenced average annual diameter growth. White pine blister rust was found in less than 2% of the trees surveyed. Signs of the red turpentine beetle were discovered in less than 2% of the population. Root and butt rot, or discoloration were found in 18% of the 161 cores examined, suggesting that decay in the stands sampled was relatively high.

In general, the white pine forest appears in a relatively healthy condition. Environmental stresses, such as ozone damage and other forms of air pollution, interacting with biological stresses still loom as a threat to these forests. The forests need to be continually monitored for survival and growth -- especially since New Hampshire is a rapidly developing state.

REFERENCES

Buchman, R.G., S.P. Pederson and N.R. Walters, 1983: A Tree Survival Model With Applications to Species of the Lakes Region. Can. J. For. Res. 13: 601-608.

Capps Jr., O., 1983: Qualitative and Censored Response Models. A.A.E.A. Econometrics Refresher Course Handbook. Aug 3-5 1983. Purdue Univ. 60+ pp.

Clutter, J.L., J.C. Fortson, L.V. Pienaar, G.H. Brister and R.L. Bailey, 1983: Timber Management: A Quantitative Approach. John Wiley and Sons. 333 pp.

Hamilton Jr., D.A. and B.M. Edwards, 1976: Modeling the Probability of Individual Tree Mortality. U.S.D.A. For. Serv. Res. Pap. INT-185. 22 pp.

Jastrzembski, J.P., 1989: Monitoring Eastern White Pine Mortality in New Hampshire. Masters Thesis. University of New Hampshire. 82pp.

Kleinbaum, D.G., L.L. Kupper, and H. Morgenstern, 1982: Epidimologic Research Principles and Quantitative Methods. Duxbury. 529pp.

National Inventories and Inventories of Endangered Forests in Europe

Michael Koehl
Swiss Federal Institute for Forest, Snow and Landscape
Dept. NFI
CH-8903 Birmensdorf, Switzerland

ABSTRACT

In Europe, large-scale forest surveys have a long history. As early as 1878 a large-scale inventory was conducted in Germany to obtain data on structure and production. Today national inventories based on mathematical and statistical methods are a matter of course in most West European countries, even though much information from inventories in numerous enterprises and some very intensive documentation of operations has already been recorded.

The basic aim of any national inventory is to obtain and furnish data to provide a basis for decisions and planning regarding forest functions. In those Central European countries with intensive forest management and dense populations, detailed data on forest area, forest condition, timber reserves, infrastructure, and the protective and recreational functions of the forest are needed.

Since the onset of forest decline in the early '80s, much emphasis has been laid on the determination of the state of health of the forest, including not only qualitative changes but also the increasing spread of this 'new type of damage'.

At this point it is appropriate to give a brief survey of national and forest damage inventories, with particular emphasis on the determination of the state of health of the forest. More detailed information on the national forest inventories of France, Italy, Hungary, Luxemburg, the Netherlands, Belgium, Spain, Portugal and Greece may be found in the publication by Pelz und Cunia (1985).

National Inventory Systems

The first national inventory in **Sweden**, conducted between 1923 and 1929, surveyed systematically distributed strips 10 m wide. Statistically, this was stratified, systematic cluster sampling. This sampling design is still being applied, though in the third national inventory, carried out in 1953, the tract system, comprising a square or rectangular arrangement of sampling areas, was introduced.

Because of great variation in forest structure, work regulations, and size and form of administrative bodies from north to south, the country was divided into 5 regions for the purposes of the 6th inventory, which started in 1983.

The size of the survey area varies; in southern Sweden one square tract contains 8 circular sample plots, spaced at 50 m. In the other regions, each tract includes 12 circular sample plots, spaced at between 100 m in the south and 150 m in the north. In order to allow continuous observation, permanent sample plots are marked out in addition. Since surveys of permanent sample plots always involve more work, whether in the first or in subsequent surveys, both the number of circular sample plots and the distance between them was reduced. This was the only way to ensure that one tract could be surveyed in one working day.

In addition to these circular plots, where the standing reserve is measured, sample plots for the measurement of annual yield and the observation of regeneration are marked out, but only temporarily, even when they are located on the permanent tracts.

In surveys of timber volume, every tree with a DBH over 10 cm is tallied. These trees are selected with probability proportional to their diameter (PPS) for estimating volume, age, and increment. Single tree volume is calculated with a volume function taking, e.g., diameter at 1.3 m and at 7 m, crown length, and tree height into account. These data are supplemented through the determination of increment, mortality, and site and stand conditions. In the temporary sample plots, increment and age are determined through sample cores.

In **Finland**, national inventories have been regularly conducted ever since 1921. The latest inventory period was 1977-1984.

In 1963, a systematic arrangement of sample tracts was introduced. Square plots were later replaced by L-shaped clusters, one of the main reasons for this change being time allocation. Especially in central Finland, it is often difficult to take continuous measurements in rectangular sample plots, as in many cases part of the tract is occupied by a lake. Comprehensive studies have shown that the survey of L-shaped clusters is more efficient.

The long sides of the L-shaped clusters measure 4100 m. Characteristics of site and stand and also yield are surveyed over 41 sampling areas 100 m apart. At 21 points 200 m apart angle point sampling is carried out with a factor of 2, the DBH and species of each tally tree being recorded.

Timber volume is determined on units of four sample plots using volume functions which include DBH, tree height, and diameter at 6 m measured with Finnish calipers as independent variables. Further, the increment on the selected plots over the past five years is determined through sample cores.

In 1984 and 1985 3,000 permanent sample plots were systematically set up over the entire country to allow predictions on production and the monitoring of ecological and environmental changes.

In the northernmost areas of Finland (Lappland), which comprise c. 9.5 million ha of forest and 5.4 million ha of managed forest, a different inventory concept is applied. Here, the basically homogeneous forest structure , the slow rate of increment, and the poor accessibility render a combination of terrestrial surveys and aerial photographs most advantageous.

In 1970 the grouping method - a special application of two-phase sampling for stratification was introduced. This method was used with only slight adaptation for the 7th inventory.

Aerial photographs are taken of strips 5 km wide and 11km apart at a scale of 1:50 000. The unit for photo interpretation is a tract of 5 x 8 km. Within each of these a systematic arrangement of sample plots allows for measurement of stand characteristics according to which the sample plots are stratified into homogeneous groups.

Terrestrial angle point sampling is then applied to about one third of these areas, selected systematically, independent of their stratification. In some of the terrestrially surveyed plots, the centre is permanently marked but not the trees assessed.

The data from terrestrial surveys are carried over to the entire aerially photographed sample to allow the deduction of results even for smaller areas independently of the latest field surveys.

In **Austria**, the national forestry inventory began in 1961, as a result of the forest survey which was conducted between 1952 and 1956 as a kind of large-scale forest management using the age class method, and which is not to be regarded as a mathematical or statistical assessment. In 1961, regularly distributed tracts were surveyed. The size of the tracts has been altered over the years. Today the entire inventory area is divided into uniform quadratic tracts, comprising four permanent angle point sample plots each and spaced at 200 m intervals.

Between the point samples, halfway along each side of the tract, sample plots are laid out on which the characteristics of the area are surveyed in order to achieve a better classification of stand characteristics. Trees sampled for volume determination are distributed over the four permanent point samples. For all selected trees with a DBH of more than 10.4 cm, the height is measured with a relascope, the diameter at one third of the tree height with a tele-relascope, and the volume is then determined by means of volume functions. Within a circular sample plot, the height and DBH of trees with a DBH between 50 and 104 mm is additionally recorded.

Up to 1980, increment was determined through sample cores, but then it was recognized that recovery from coring may influence growth in diameter and coring should consequently not be employed on permanent sample plots. Since then, one of the sample plots has been selected in each case for special scrutiny: in the first subsequent survey, the normal parameters will be measured, but core samples will be taken as well, so that the results from the two methods can be compared (core vs. comparison of mass of individual sample stems on the permanent sample plots at different points in time).

The tracts are distributed at regular intervals of 2.75 km over the entire country. Each year, some 1100 tracts, also equally distributed over the entire country, are surveyed. As a result, the sampling network was surveyed in a regular pattern between 1981 and 1985. In 1986 the first checks on permanent sample plots were made. At the same time surveys are being conducted on temporary sample plots.

In 1978 the Federal Ministry for Nutrition, Agriculture, and Forestry commissioned a specialist committee with the task of elaborating a national forest inventory for the **Federal Republic of Germany** based on sample plots. Field surveys were conducted between 1986 and 1988, and the results are expected to appear by the end of 1989.

Sample plots are regularly distributed over the entire country in a 4 km grid. In some states or areas, it has been possible to condense the network to a 2.83 km or even 2 km grid.

Square clusters with 600 m long sides centred on the intersection points of the grid are surveyed. The sides of the tracts serve as assessment lines for the determination of stand boundaries, stand characteristics, and an inventory of roads and paths.

At each corner of the tract sample stems (DBH >10 cm) are assessed through angle point sampling, using a factor of 4, and young growth is surveyed over a concentric plot (r = 1,2, or 4 m).

Where the centre of an angle point sampling plot lies close to the limit of a forest area or stand, the mirror image of the centre is located and sample trees distributed around it are selected.

For each sample stem the location is recorded from the azimuth and horizontal distance, and the species, age, a given stem index, social status within the stand, and damage are documented. As well as DBH and overall height, diameter at 7 m is measured with either Finnish calipers or a relascope.

The first national forest inventory in **Norway** was conducted between 1919 and 1932. Field surveys for the sixth survey, conceived as a one-phase, systematic sampling procedure, started in 1987.

Sample plots are distributed in a 3 km grid over 13 counties, permanently marked, and surveyed at 7-year intervals.

In contrast to earlier surveys in which relascope points were employed, the latest surveys use concentric circular plots. In the smaller plots of radius 5.65 m (100 m2), all trees with a DBH of 5 cm or more are registered, and in the larger plots of radius 7.98m (200 m2) all trees with a DBH of at least 20 cm.

Around these permanent circular plots temporary plots are laid out, their number and spacing depending on the size, distribution of forested areas, and forest characteristics of each county.

The basic idea of a comprehensive national inventory for **Switzerland** was first mooted by A. Kurth in 1956. In 1973 a special department, 'National Forest Inventory' was set up at the Swiss Federal Institute of Forestry Research, and commissioned with the planning and execution of a national forest inventory. After preliminary trials in the canton of Nidwalden, terrestrial surveys were conducted between 1983 and 1985.

As a first step, data on the area and stand, such as ownership, exposure, erosion, regeneration, potential yield, etc., are registered. Sample trees are selected from two concentric plots beeing systematically distributed over a 1 km grid and covering 0.05 ha and 0.02 ha. On the smaller plots all trees with a DBH over 12 cm are surveyed, on the larger ones all those with a DBH over 36 cm, with DBH, species, storey, crown class and degree of stem damage being recorded.

All trees with a DBH over 60 cm and an azimuth of 150g or less towards north are taken as tariff trees; here the height and diameter at 7 m area also measured, so as to allow volume determination by means of volume functions $V = f(d1.3,d7,h)$. The upper diameter is measured with Finnish calipers.

Data on young growth, i.e., from trees with a DBH of less than 12cm, are collected on circular sample plots with a radius of 3 m. In order to reduce the amount of work, 30 plants are selected from those with a DBH of less than 4 cm or less than 1.3 m high, beginning with an azimuth of 0g (= north), the azimuth of the 30th plant determined, and the number of plants in the whole plot extrapolated. Further, data on stand stability are collected on every sample plot.

In addition to the terrestrial survey, a questionnaire is circulated among district Forest Offices to obtain data on ownership, yield, and accessibility.

Forest Damage Inventories

Terrestrial Inventory Procedures. The main criteria employed in assessing the 'new type of forest damage' are foliage loss, discoloration, and degree of discoloration. According to these criteria, trees may be allocated to damage classes. To differentiate between this type of damage and that obviously due to other causes, further information is needed, such as data on the individual tree, clear sources of damage, and a general description of the state of health (see Tab. 1). The major sources of error in visual assessment are listed in Tab. 2.

The accuracy of assessment depends on the ability of the surveyor to classify the tree correctly. The density or sparseness of a tree crown is determined by many phenomena acting in certain sections or even on particular branches. In addition to this complexity, individual traits deriving from heredity, site conditions, or social status come into play. Consequently, determination of foliage loss is very tricky. Some authors believe that differences in assessment can be obviated by thorough and uniform training of field crews, so that the only errors are random in nature, but others question this.

Table 1. Characteristics recorded in terrestrial damage inventories.

Site characteristics	elevation, exposure, relief, inclination, soil conditions;
Data on individual trees	social status crown form, crown length exposure, influences due to position near margin crown visibility flowering effect in pine;
State of health of the trees	injury to stem or bark stem canker bark necrosis stem fracture frost rib, mistletoe anastomoses, tortuous growth resin exudation game damage
Obvious cause of damage	harvesting insects, fungi, game fire, erosion, landslips rockfalls snow, wind, frost, lightning, sun
Features not attributable to known causes	foliage loss foliage discolouration epicormic shoots and clusters intensity of fructification top dying

Table 2. Sources of error in damage assessment

1. Sampling error

2. Non-sampling errors
   - random error
     .error in estimation
   - systematic errors
     .influence of weather
     .determination of a "healthy" reference tree
     .determination of the exact beginning of crown
     .estimation of flowering effect in pine
     .treatment of regenerated crowns
     .consideration of phenotype
     .evaluation of different leaf sizes
     .individual errors among crews
     .delimitation of new type of forest damage from other causes

In Sweden, Westmann and Lesinski (1985) investigated the accuracy with which needle loss is assessed by comparing the assessments of field crews with the loss in biomass of representative shoots. They found that the assessments, in which trees were classified in 10% intervals, were correct for 22% of the trees. For 70% of the trees surveyed, crown thinning had been underestimated by one or two classes. Great errors mainly occurred in regard to slightly damaged trees, while trees with very sparse crowns tended to be estimated more exactly. Such differences in estimation render continuous control essential. Deviations from the norm may be controlled through contiuously switching field crews and intensive training even during surveys. Bias in the estimation of spatial distribution introduced by field crews is avoided by shifting the groups from one area to another.

Another problem is the choice of the reference tree. In most cases, nobody knows what the perfect healthy tree looks like. In some countries (e.g., Switzerland, Federal Republic of Germany) field crews are provided with colour photographs to help them in diagnosing and classifying forest damage.

The assessment of forest damage requires a representatively distributed network of permanent sample plots, over which consecutive surveys can be conducted. Such networks have now been constructed for most countries in Europe. Examples of large-scale forest damage inventories (as of 1988) are given in Tab. 3b, showing the inventory methods of 16 out of the 35 parties participating in the "International Cooperative Programme on Assessment and Monitoring of Air Pollution Effects on Forests".

IRC Inventories. Infra-red aerial photographs were used for the determination of forest damage as early as the mid-'60s. Differences in reflections from different plant species and plants with varying degrees of injury are particularly apparent in the red, green, and infra-red parts of the spectrum. Consequently, films especially sensitive in this range are employed in forest damage inventories. The aim is not to achieve a true reproduction of the natural object but to emphasize contrasts as much as possible, so as to be able to differentiate as finely as possible. Tissue degeneration reduces reflection in the infra-red zone even at an early stage of decay, before noticeable changes in the visible range of the spectrum become evident.

As infra-red color aerial photography is much in use in forest damage inventories, the reader is refered to Schmid-Haas (1985) and Koehl (1987a, 1987b) for further information.

Critical Remarks on Forest Damage Inventories. Several points on which current forest damage inventories have been criticized can be compiled as follows:

- studies on vitality do not adequately differentiate between entire stands and individual trees;
- while inventories in some countries cover only hard-hit areas, those in other countries cover a nationwide sampling grid;
- differences in the criteria used to describe forest damage and in the way in which the results are compiled make comparisons difficult;
- in some cases, too few variables are determined:
- the extrapolation of values from individual trees for the determination of data on particular areas is not really valid (percentages of damaged trees do not necessarily reflect percentages of areas with damage);
- the question of reference trees has not been satisfactorily settled;
- the separation of known and unknown reasons for foliage loss is done in a subjective way;
- the relation between vitality and foliage loss is not satisfactorily verified.

These points impose the conclusion that results from inventories conducted in different countries and/or in different years are difficult to compare. Because the need for an international adjustment is obvious, the ECE worked out guidelines for Forest Damage Surveys. In 1988 most European countries followed these guidelines in assessing forest decline.

Tab.3a Assessment Methods According to ECE-Guidelines

Defoliaton			Discoloration		
class	needle/leaf loss	degree of defoliation	class	foliage discolored	degree of discoloration
0	up to 10%	none	0	up to 10%	none
1	11-25%	slight	1	>10-25%	slight
2	26-60%	moderate	2	>26-60%	moderate
3	>60%	severe	3	>60%	severe
4	100%	dead tree			

SUMMARY

National forest inventories in Europe are characterized by the transition from a pure determination of current conditions to the monitoring of changes. The aim of such monitoring is not only to observe changes in in productive function of the forest; in forest damage inventories the entire forest is considered as an ecosystem.

No matter what the difficulties, the estimation of leaf or needle loss still remains the only possible means of determining the vitality of the trees.

Permanent monitoring of the forests is necessary to decide whether the foliage loss observed from the beginning of the present decade indicates a natural, short-term fluctuation or a long-term trend.

LITERATURE

Cruse T. et al.,1985: Designing a new national survey for Sweden,Swedish University of Agricultural Sciences,Umea.

FAO,ECE,Finnida,1987: Ad hoc FAO/ECE/Finnida Meeting of Experts on Forest Resource Assessment, Proceedings, Bulletins of the Finnish Forest Research Institute Nr.284.

Koehl M., 1987a: Forstliche Nationalinventuren in Europa,AFZ,No.1/2, pp.19-24.

Koehl M., 1987b: Gegenwärtiger Stand der Methoden zur Inventur und Ueberwachung gefährdeter Wälder, AFZ, No.22, pp.573-577.

Pelz D.R. and Cunia T. (ed.),1985: National Forest Inventoriy in Europe, Procedings,Mitteilungen der Abteilung Forstliche Biometrie der Universität Freiburg,Nr.1.

Schmid-Haas P. (ed.),1985: Inventorying and Monitoring Endangered Forests, IUFRO Conference,Zurich,Switzerland.

Westmann L. and Lesinski J.A., 1985: in: Inventorying and Monitoring Endangered Forests, IUFRO Conference,Zurich,Switzerland.

Tab. 3b Forest Damage Inventories in Europe (1988)

Country	sampling design	assessment of defol./disc.	remarks
ECE-Guidelines	16X16km grid, fixed area plot or cluster of trees	according to tab.3a	
Austria	4X4km,4X2km,2X2km	in 5 classes	only stands >60 years
Belgium		according to ECE	
-Flandern	8X8km,24trees/plot		5 extra plots
-Wallonia	16X16km grid		
Denmark	-global assessment of stands -16X16km grid	2 classes according to ECE	
Finland	-systematically distributed plots -16X16km grid	disfoliation in 10%-classes	part of the national forest inventory higher density in south
France	16x16km grid, 24 trees/plot	according to ECE	densitiy changed from 16X1km to 16X16km
Germany (FRG)	4X4km,8X12km 24trees/plot	according to ECE	
Greece	16X16km grid	according to ECE	all high forests throughout Greece
Hungary	4X4km grid	national classification, transformed to ECE	1st representative survey in 1987
Italy	16x16km,4X4km, 2X2km(Tuscany)	according to ECE	3X3km grid of NFI
Luxemburg	2X2km grid	according to ECE	additionally 12 plots in spruce stands
Netherland	1X1km grid	according to ECE	
Norway	9x18km grid, 1plot(250m^2)	according to ECE	in productive forests
Portugal	16X16km grid		1st inventory in 1988
Spain	16X16km grid, 24 trees/plot	according to ECE	
Sweden	SPR,systematic sampling, variing grid network,5-10 plots/cluster	disfoliation assessed in 10%-classes, class-widths of 20% for national records	permanent and temporary plots, remeasurement begins in 1989
Switzerland	4X4km grid 500m^2 plot	according to ECE estimates in 5%-classes	1X1km grid of NFI

Indicator Plants in Forest Health Surveys

Paul C. Van Deusen and Joseph E. Barnard
USDA Forest Service Mathematical Statistician, 701 Loyola Ave., New Orleans,
LA 70113, and USDA Forest Service, Program Manager of the National Vegetation
Survey, Research Triangle Park, NC 27709

--ABSTRACT--

The USDA Forest Service (USFS) moved toward implementation of a National Forest Monitoring system by conducting a Visual Damage Survey (VDS) during the 1988 field season in selected Eastern states. The VDS is designed to be compatible with similar surveys being conducted in many European countries. At this point, some of the more interesting results are related to the use of indicator plants that are sensitive to the presence of certain pollutants. This paper presents results based on the 1988 VDS data and discusses the implications of these findings for future work and research related to indicator species. Our preliminary analyses suggest that indicator plants are providing information that is correlated with individual tree growth rates.

INTRODUCTION

Concerns over atmospheric deposition influences on forests (ADIF) in the U.S. were awakened by the observance of wide spread forest damage in a number of European countries. The causes of this forest damage are still a matter for debate, but factors such as acid rain and photochemical oxidants are often considered. An International Cooperative Program on Assessment and Monitoring of Air Pollution Effects on Forests was launched late in 1985 with the assistance of the United Nations Environment Program (UNEP) and the United Nations Economic Commission for Europe (UN-ECE). Although the U.S. is a member country, the first forest condition survey was not made until 1988, and we will refer to this as the visual damage survey (VDS).

The VDS is likely to be a recurrent activity for several years to come and will generate numerous problems of interest to forest researchers. The VDS should evolve into a formal forest monitoring activity currently being planned by the USFS. It is expected that Environmental Protection Agency (EPA) monitoring activities for forests currently being planned under EMAP (Environmental Monitoring and Assessment Program) will also link up with USFS monitoring. Considering the various monitoring activities that are either ongoing or envisioned along with the mandate recently given to the USFS by the Forest Ecosystems and Atmospheric Research Act of 1988, it is evident that some form of forest health monitoring and research will be funded for years to come.

The National Vegetation Survey has responsibility for funding and organizing the planning and field activities associated with the VDS. The field samples are taken at permanent Forest Service Inventory and Analysis plots (FIA) and at Forest Service Forest Pest Management plots; so the additional measurements taken for the VDS can be linked with existing USFS data and field activities. Measurements of stand and tree characteristics include assessment of indicator plants for air pollution, various evaluations of the foliage, and root samples are taken to unearth diseases. The details of these measurements can be found in the VDS field manual (Alexander and Carlson 1988). The majority of the indicator plant data for the 1988 VDS pertain to ozone, and we will discuss only the analysis related to ozone for the Southeast.

POLLUTANTS AND INDICATOR PLANTS

Three major pollutants have been demonstrated to cause significant foliar injury to forest trees (Skelly et. al. 1987): ozone, sulfur dioxide, and fluoride. Other pollutants have been shown to have local effects, such as hydrogen chloride, bromine, and nitrogen oxides; but these are of limited use for forest monitoring. Ozone (O_3) is very widespread in the U.S. and is believed to cause more plant injury than any other air pollutant. Ozone is a secondary pollutant formed by photochemical reaction of hydrocarbons and oxides of nitrogen. Natural sources of ozone do not contribute significantly to ground level concentrations (Skelly et. al. 1987).

The visual symptoms of ozone exposure are generally recognized as either acute or chronic. Acute injury develops within a few hours or days after exposure and is expressed as stippling, flecking, bleaching, and bifacial necrosis. Acute injury is usually in response to high ozone concentrations, while chronic injury results from long-term low-concentrations. Chronic injury may be manifested as chlorosis, stippling, premature senescence, and necrosis.

Tree species that are known to be bioindicators of ozone include: white ash (*Fraxinus americana* L.), black cherry (*Prunus serotina* Ehrh.), flowering dogwood (*Cornus florida* L.), eastern white pine (*Pinus strobus* L.), sassafras (*Sassafras albidum* (Nutt.)), and yellow-poplar (*Liriodendron tulipifera* L.). See Skelly et. al. (1987) for excellent photographs of the foliar symptoms on these species and more. The VDS used black cherry and yellow poplar as the bioindicators for the Southeast. Characteristic visual symptoms of Ozone were recorded as either present (level 1) or absent (level 0). It is generally believed that the symptoms caused by ozone to designated indicator plants are unique and can be recognized with little error (Samuel A. Alexander pers. comm. 1989).

ANALYSIS OF THE SOUTHEASTERN DATA

A limited number of sites were visited in Georgia, North Carolina, and South Carolina during the 1988 field season. The basic field measurements were made at FIA plots, which are actually clusters of horizontal point samples. The explanatory variables of interest here are the indicators of ozone (O3) that came from either black cherry or tulip poplar, DBH at the 1988 measurement, and crown-ratio in 1988 (CR). We will look at 3 response variables: 1) 5 year radial increment (R5) from tree cores, 2) crown discoloration (CD) as ocularly estimated to the nearest 5 percent class, and 3) defoliation (DEFOL) as ocularly estimated into 5 percent classes. The basic analysis will be to fit a linear model with O3 treated as a classification variable and all others treated as continuous. The adjusted means of the response variable for the 2 values of ozone will then be statistically compared to test for an effect when ozone was indicated at the plot. Least squares adjusted means have been 'adjusted' for the effect of DBH and CR so the comparisons are hopefully free from the inherent ability of trees to grow differently based on diameter and crown size.

In Georgia, 96 plots were measured and 28 indicated ozone; in North Carolina, 14 plots were visited and 7 indicated ozone; while in South Carolina there were 48 plots with 14 indicating ozone damage. There is not enough data to be meaningful in North Carolina, but we present the results for all 3 states to be consistent. Additional data will be collected in North Carolina in the 1989 field season to remedy this. The analysis only includes loblolly pine (*Pinus taeda* L.) with all variables non-missing -- for the R5 analysis there were respectively 307, 43, and 155 trees for Georgia, North Carolina, and South Carolina; for the CD analysis there were 782, 158, and 384 trees; and for the DEFOL analysis there were 782, 158, and 384 trees. Thus, the weakness of the North Carolina results is again evident.

RESULTS OF THE ANALYSIS

The data for each state were fit to the following linear model:

$$[1] \qquad Y_{ij} = a_j + bD^2_{ij} + cCR_{ij} \, ,$$

where Y is the response variable, D^2 is DBH squared, and the parameters to be estimated are a_j, b, and c. The *j* subscript is used to denote the 2 levels of O3, and *i* denotes tree i. Thus, the model implies that both levels of O3 have the same slopes for D^2 and CR, and only the intercept (a_j) changes when O3 is indicated.

Table 1 gives the results for each state for the 3 response variables that were fit to model [1] independently, i.e. MANOVA was not used. A parameter can be judged significant at approximately the 95% level if the absolute value of the T-ratio is 2 or more. Thus, most of the parameters are significant for Georgia and South Carolina, while many are not significant for North Carolina, probably due to the small sample size.

TABLE 1. Results of fitting model [1] to the loblolly pine data.

Response Variable	Parameter		Georgia Estimate	T-ratio	North Carolina Estimate	T-ratio	South Carolina Estimate	T-ratio
R5	O3*	0	3.99	3.42	-1.87	-0.42	4.77	2.30
		1	3.38	2.97	-3.31	-0.94	2.27	0.99
	D^2		0.0088	3.48	-0.0085	-1.21	0.0058	2.38
	CR		0.101	3.41	0.332	3.00	0.115	2.29
DEFOL	O3	0	32.66	18.52	39.86	8.92	47.96	13.47
		1	33.95	20.05	38.46	9.83	51.52	13.41
	D^2		-0.01	-3.75	-0.03	-2.95	-0.03	-5.72
	CR		-0.37	-8.31	-0.42	-3.63	-0.54	-6.28
CD	O3	0	12.26	8.29	8.69	3.69	14.24	7.38
		1	13.16	9.26	7.73	3.74	14.15	6.69
	D^2		0.010	3.24	0.002	0.32	-0.002	-0.96
	CR		-0.189	-4.99	0.022	0.35	-0.087	-1.88

*/ O3 denotes the intercept in model [1].

Table 2 gives the results of testing for differences in the least squares means (LSmeans) of the response variable. This comparison tests for differences in the mean of the response variable when O3 is present after adjusting for the effects of D^2 and CR. The p-value in Table 2 results from a 2-sided t-test where the null hypothesis (H0) is that the LSmeans are equal regardless of the value for O3. The p-value is very conservative, because it would also be reasonable to test H0: growth is less when O3=1, defoliation is greater when O3=1, and crown discoloration is greater when O3=1. If these 1-sided tests were made, all of the above p-values would be divided in half and the results would generally appear more significant. Of course, in the cases where the wrong relationship exists -- like with defoliation being less when O3=1 for North Carolina -- the 1 sided null hypothesis cannot be considered at any p-value. The most meaningful way to evaluate Table 2 is to look for similar trends across the 3 states, keeping in mind that there is little data for North Carolina, and to use the p-values as an indication of how seriously to take the results.

Table 2. Testing for differences in least squares means for the 2 levels of ozone using model [1] and the loblolly pine data.

Response Variable	O3 level	Georgia LSmean	Georgia P-value	North Carolina LSmean	North Carolina P-value	South Carolina LSmean	South Carolina P-value
R5	0	9.58	0.238	9.64	0.386	10.98	0.005
	1	8.97		8.20		8.48	
DEFOL	0	15.12	0.096	18.36	0.408	19.64	0.015
	1	16.40		16.97		23.20	
CD	0	6.66	0.162	9.83	0.660	10.04	0.913
	1	7.55		8.86		9.95	

The radial increment results have the most direct economic and biological interpretation, and in Table 2 the LSmeans are consistently less when O3 = 1. Only the South Carolina results would be judged significant by the p-value, but the consistency across the 3 states suggests that these results are not spurious. The LSmeans for DEFOL are significant and in the anticipated relationship for Georgia and South Carolina. One would expect that exposure to O3 could cause the tree's foliage to be less full. The LSmeans results for CD are equivocal; for Georgia, the CD results are as one might expect, but for the other states they show an unexpected and weak relationship with O3.

CONCLUSIONS

This study is of interest because it suggests that ocular estimates of the presence or absence of ozone damage are correlated with radial increment for loblolly pine in the Southeastern United States. The analysis also suggested that defoliation was greater when ozone was indicated. However, one cannot assume that ozone was the direct cause of effects such as reduced individual tree growth. A number of other variables not considered in this analysis could have caused these effects as well.

When dealing with subjective measures such as ocular estimates of defoliation and ozone, one might be concerned that the field crews would not make their assessments in a completely independent manner. Specifically, they might tend to estimate a higher level of defoliation after deciding that ozone is present. However, it is difficult to imagine that the field crews could be similarly affected when they core the trees. Thus, the relationship of radial increment and the O3 indicator along with the defoliation results suggests that indicator plants can play an important role in forest health assessments.

LITERATURE CITED

Alexander, S. A. and J. A. Carlson. 1988. Visual damage survey pilot test project manual. Available from: Forest Pathology Lab, Glade Rd., Virginia Polytechnic Institute and State University, Backsburg, VA. 24061-0330.

Skelly, J. M., D. D. Davis, W. Merrill, E. A. Cameron, H. D. Brown, D. B. Drummond, and L. S. Dochinger, Eds. 1987. A manual for diagnosing injury to Eastern forest trees: air pollutants, pathogens, insects, and abiotic stresses. The Pennsylvania State University, College of Agriculture, University Park, PA 16802.

Spacial Analysis of Inventory Data

Moderator: Joseph C. Mawson

Carlton M. Newton and Steven T. Bower, Spatial Analysis of Forest Inventory Data

William H. Mc Williams and Joanne L. Faulkner, Spatial Analysis Using a Systematic Grid of Forest Inventory Plots: Selected Examples for the South Central United States

Robin M. Reich, Paul W. Mielke, Jr., and Frank G. Hawksworth, Application of Permutation Techniques in Analyzing Spatial Patterns

Spatial Analysis of Forest Inventory Data

Carlton M. Newton and Steven T. Bower
Associate Professor and Graduate Student
School of Natural Resources
University of Vermont
Burlington, Vermont 05405

ABSTRACT

Concepts basic to the spatial analysis of forest inventory data are introduced, with special attention to autocorrelation and the estimation of population totals and means. In most cases, the estimates based on spatial analysis procedures are not expected to equal those resulting from traditional statistical estimation procedures. Initial findings of a comparative study are briefly discussed.

INTRODUCTION

The primary reason for this paper at a workshop on state of the art methodology is the relatively widespread availability and use of geographic information systems (GIS). A GIS is a combination of computer hardware, software, geographically referenced data, and trained personnel that collectively facilitate the storage, retrieval, and manipulation of the data according to either their spatial locations or their associated attribute values.

Because of the ease of entry, the forest inventory data layer is likely to be one of the first to be created in the GIS data base. The attribute portion of the forest inventory data is probably already computerized and the sample units are easily entered as geographic point locations in a two-dimensional space, i.e. spatially referenced with a single pair of (x,y) coordinates per sample unit. While it was probably not anticipated at the time of collection, the forest inventory data can now be subjected to several analysis techniques that are common in a GIS. Generally referred to as spatial analysis techniques, these methods have historically been more frequently associated with the field of quantitative geography than with forestry.

Our intent is to briefly introduce and discuss some basic concepts, analysis techniques and algorithms that are used to analyze a point data layer. The techniques and algorithms that we have selected are those that we believe are likely to be applied to forest inventory data in a GIS environment. At present, we are in the process of completing some research comparing area estimates generated from these techniques with those obtained from traditional statistical methods. While we will present some of our initial findings, time and space prevent us from discussing in depth the individual or comparative error assessment techniques.

SPATIAL ANALYSIS CONCEPTS

Spatial analysis is the analysis of geographically referenced data for the purpose of identifying and describing spatial relationships and patterns. Most introductory texts in quantitative geography present the general underlying definitions and assumptions of spatial analysis techniques (for example see Amedeo and Golledge, 1975, Taylor, 1977 or Unwin, 1981). A few of these concepts seem particularly relevant to the general spatial analysis of forest inventory data.

Data Characteristics

The sample units in a forest inventory, even if they are fixed radius plots, are usually considered to be points located in the two-dimensional space of the area inventoried. The exception of course is where the sample units were plots that were rather large relative to the area inventoried -- in which case the plots would be considered polygons. Therefore, for a given variable, inventory data are most often represented as a set of points in a three-dimensional space. The location of an individual sample is given as a pair of (x,y) coordinates on a rectangular reference system and the observation is represented by a third-dimensional z value.

The population is then envisioned as a surface from which a sample is drawn. The rougher, or more spatially variable, this population surface is, the more difficult it is to model, or estimate, for a given sample size. Similarly, the smoother or more uniform the population surface, the easier it is to estimate. Different sample designs are characterized by a different set of (x,y) values, accompanied presumably by different z values. A remeasurement of permanent plots produces a data set that consists of two z values for each (x,y) -- one z value for the first measurement and the second for the remeasurement.

One objective of the inventory could be the estimation of the true but unknown population surface. Other samples, that is different (x,y,z) data sets, or different estimating algorithms are of course expected to yield different estimated surfaces. Although calculated somewhat differently, the concepts of precision, accuracy, and bias are as applicable to surfaces as they are to traditional descriptive statistics. Hsu and Robinson (1970), MacEachren and Davidson (1987), Davis (1976) and Shepard (1984) all provide methods for assessing the differences between an estimated surface and the population surface.

Spatial Autocorrelation

When dealing with spatial data, it is important to understand the concept of spatial autocorrelation. If the value of a variable at one observation is not totally independent of a value at an adjacent or nearby observation, then spatial autocorrelation exists. As we pass through a forest we often see examples of spatial autocorrelation in terms of slow or gradual changes in variables such as stems per acre, mean canopy height and basal area per acre. The major problem with spatial autocorrelation is that because the variates are not independently distributed, a bias is introduced in several traditional statistical estimates. Within the field of forest inventory, Matern (1947) was the first to acknowledge autocorrelation and its impact on forest inventory. Matern has been one of the few forest researchers who has directly used spatial autocorrelation as a guiding principle in inventory design (Matern, 1960).

From the perspective of the population surface, spatial autocorrelation is a measure of surface variability, or roughness. By themselves, the attribute or z values from a sample allow the calculation of traditional estimates of variability. However, differentiation among surfaces that appear to be rugged mountains, rolling hills, and nearly uniform plains, requires the evaluation of relationships between the z values of neighboring locations. As one traverses the surface, a plain would be characterized by a linear change of z values. For rolling hills a z value of a chosen point will tend to be more similar to the z values that are close to it than those that are more distant. A rugged surface would have z values that change radically, upwards and downwards, as one moves away from a given point.

There are only a few techniques that are commonly use to assess autocorrelation in a sample. The correlogram is the curve or function of correlation against distance that is derived by plotting the set of correlations for all pairs of z values over the distance separating the observations (Cochran, 1977). In addition to the use of a correlogram, one can assess autocorrelation with Moran's I statistic

(Moran, 1950; Cliff and Ord, 1981). For this statistic, the correlation for adjacent data is compared to that which would be expected for a random distribution. Moran's I can assume values from -1.0 to +1.0. Maximum position autocorrelations will give an expected value of I equal to +1.0. A value close to zero is expected from a random distribution of values. A negative value indicates negative autocorrelation which arises when neighboring z values are distinctly dissimilar. Geary's c statistic, on the other hand, has values that range from zero to +1.0. A c value close to zero indicates that there is little variation in the z values, thus implying a strong autocorrelation. A value close to 1.0 indicates a near random distribution, and little or no autocorrelation (Geary, 1954).

Both Moran's I and Geary's c can be distance weighted to reflect the relative influence of the separations for different pairs of z values. Cliff and Ord (1981) have determined that the variance of either statistic can be calculated and used in tests of significance. They have also shown that tests based on Moran's I are consistently more powerful than those based on Geary's c. A good, concise introduction to spatial autocorrelation is given by Odland (1988).

Spatial Interpolation and Extrapolation

The process of deriving a map, irrespective of the type, from sample data involves spatial interpolation and quite often extrapolation. Simply put, spatial interpolation is the process of estimating or filling in the z values at unsampled sites within the area covered by existing sample observations. There are numerous techniques and algorithms for interpolation, a few of which will be discussed later in this paper.

The minimum sized convex polygon that encompasses all of the sample points is termed the convex hull. A convex hull will fully include the entire population area if the sampling is purposely done along, or beyond, the boundary of the population area. However, this is not the case with most sampling designs. The process of extending the map beyond the convex hull, outward to the population border (or further) is called spatial extrapolation. While there are numerous interpolation methods in use, special techniques for extrapolation have not been as well developed and tested.

Total and Mean Estimates

Given a surface or map for a particular variable, how are estimates of the population total and mean derived? If a variable is on either the interval or ratio scales of measurement, these estimates can be derived from the estimated surface. The total is defined as the integral, or volume, of the geometric solid below the surface. For example, the estimate of total board foot volume of softwood sawtimber in the population is made by first estimating the map or surface for softwood sawtimber and then calculating the total volume of the solid between the estimated board foot volume surface, and the $z = 0$ plane.

Spatially, the estimate of the total is a function of the generated map or surface. Therefore this estimate will not necessarily equal the same value as that from traditional statistical methods. The estimate of the mean is calculated by dividing the estimate of the total by the population area. If the variable of interest is expressed on a per unit area basis, e.g., number of trees per hectare, then the mean will also be on a per unit area basis.

The critical issue, of course, is that estimates of the total and mean first require the derivation of the estimated surface. This approach to estimating the population total and mean ensures compatibility among the estimates, the maps or surfaces, and other spatial analysis results.

SELECTED ANALYSIS TECHNIQUES

Given the above basic concepts, there are many spatial analyses that can be applied to forest inventory data. We will, however, limit ourselves to only a few techniques that are applicable to the problems of spatial interpolation and extrapolation. Introductory coverage of a wider array of techniques can be found in texts by Pielou (1977), Upton and Fingleton (1985), Taylor (1977) and Johnston (1980).

Recall that the objective of spatial interpolation and extrapolation is the estimation of a map that is derived by first estimating values intermediate to the sample points and then, if necessary, applying an extrapolation algorithm to extend the map outward to the limits of the population space. We will only present those techniques with which we have experience and which we believe can or will be found in a typical GIS software environment. Some of our coverage of interpolation techniques follows that presented by Burrough (1986).

Interpolation Techniques

The methods available for spatial interpolation can be divided into two types -- discrete and continuous. With discrete techniques a boundary is drawn around spaces that seem reasonably uniform relative to some variable or attribute. The continuous methods assume that there is a gradual change in the variable and that some type of smooth mathematically based technique can be used to model the changes.

For discrete models, no mathematical computations are necessary to locate the boundaries, the mapped data of interest can be measured according to all scales of measurement. However, if mathematical computations are involved, then only interval and ratio data can be used. Averack and Goodchild (1984) present a concise comparison of the algorithms commonly used to define boundaries from the data. Maps of soils, stand types, insect infestations, commercial timberland, and stocking levels are examples of the results of such techniques. Analytically such maps imply (1) that within the boundaries of an area the attribute or variable of interest is homogeneous and the same in all directions, and (2) that all important variation occurs at the boundaries. Stepped models and Thiessen polygons are two techniques in this category.

The continuous models include population level techniques such as trend surfaces and Fourier series, and localized or regional techniques such as splinning, Kriging, and linear and non-linear interpretation of triangular irregular networks. For the population level techniques the model is constructed from all of the observations of the variable of interest. Therefore, these techniques don't handle local variation very well and are more applicable to problems associated with variation over large areas. The localized or regional techniques are founded on the concept of estimating individual values as a function of only the nearest neighboring observations. A window or region of sample data inclusion is positioned over each location where a z value is to be estimated. Local variation is thus accommodated at the possible consequence of not being very effective at detecting longer range trends and variations. All of these techniques require that the variable be on the interval or ratio scales of measurement.

We will limit the rest of our discussion of interpolation techniques to the step models, Thiessen polygons, trend surface functions, and the interpretation of triangular irregular networks.

Step Models. As previously mentioned, the step model is obtained by drawing boundaries around areas that are assumed to have the same z value or attribute. Figure 1 presents a three dimensional perspective of this model.

Occasionally the boundaries and the z values reflect two different variables. Such would be the

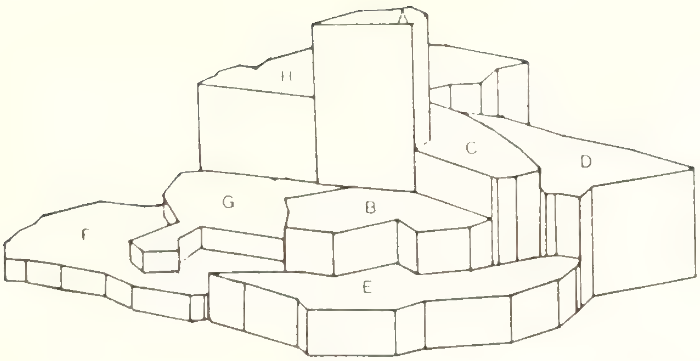


Figure 1. Three-deminsional perspective of a step model (from Burrough 1986).

case if the boundaries were for stands, and a stand's z value was the estimated mean sawtimber cubic volume per unit area. Since the geometric volume for any one of the stands equals the surface area times the solid's height, figure 1 is a visual example of what the results of a traditional stratified sample might look like. For this case, the spatial and statistical estimates of the total and mean are identical.

Thiessen Polygons. Rather than having the analyst draw a boundary around subgroups of the sample, the spatial distribution of sample units can be used to create the bounded areas. One bounded area, or polygon, is created for each sample unit, and the assigned z value comes from the sample unit. Each Thiessen polygon (also called a Dirichlet cell or Voronoi diagram) delineates all of the area that is closer to the associated sample unit than to any other sample unit (see figure 2).

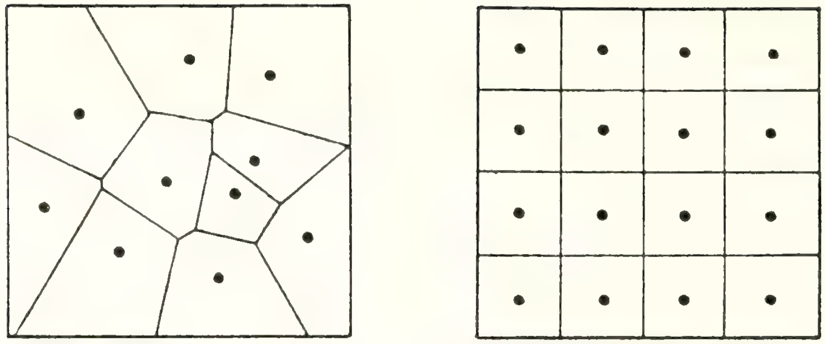
This technique for partitioning an area has a variety of uses including problems of proximity and reachability, the study of ecological distributions of plant communities (Pielou, 1977), and possibly the test for spatial randomness of the sample units (e.g., see Upton and Fingleton, 1985).

As demonstrated in figure 2, it is the spatial distribution of sample units and not the measured z values that determines the sizes and shapes of polygons. If there is a systematic distribution of sample units, the polygons become equal sized grid cells. In this special case the spatial and statistical estimates for the population total and mean are identical. Differences in the estimates are created as the distribution of sample units shifts away from systematic.

Trend Surfaces. Step models and Thiessen polygons are the most frequently used discrete models of interpolation. Of the continuous models, trend surface models are some of the most common.

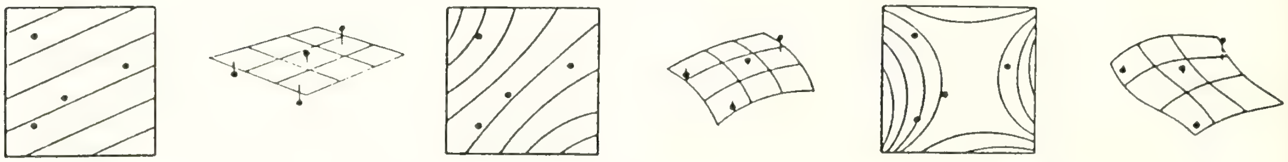
By assuming that the (x,y) coordinates are the independent variables and the z value is the dependent variable, a polynomial least squares surface is fit to the sample data. Since all of the sample data throughout the population are used to estimate the coefficients, this technique may be termed a population level model. Figure 3 portrays the trend surfaces for first, second and third order equations.

The process of creating increasingly complex trend surfaces often is approached much like a stepwise regression problem. Variables are successively added until there is no statistical significance between the current function and the next. Thus, with rougher surfaces, higher order equations are developed. A problem that arises with higher order surfaces is that the partial regression coefficients become difficult to interpret. It is because of this interpretational problem that a common use of trend surfaces is to detect spatial patterns in the residuals. Furthermore, the higher order surfaces are prone to giving unreasonably high or low extrapolations in the area just beyond the convex hull.



(a) Random distribution (b) Systematic distribution

Figure 2. Examples of Thiessen polygons for two different distributions of sample units (from Burrough 1986).



(a) Linear (b) Quadratic (c) Cubic

Figure 3. Trend surface examples (from Burrough 1986).

With trend surfaces, estimates of the population total and mean are the result of integration or numerical methods. If all three of the following requirements are met, the trend surface and traditional statistical estimates will be identical: (1) the sample must be systematic; (2) the population surface must be a plane; and (3) there must be a perfect fit of the surface to all of the sample observations.

Insofar as alternative sampling designs are concerned, Johnston (1980) advises that in an attempt to reduce the undesirable consequences of possible collinearity, systematic sampling would be the preferred design when using trend surfaces.

Triangular Irregular Networks. A triangular irregular network (TIN) was described by Peucker et al (1978) to be a very effective terrain model. Clearly, it is only a small conceptual step to go from modelling terrain to modelling other surfaces.

In a TIN, sample points are connected by straight lines to form a network of continuous connected triangles (see figure 4). Each triangle is therefore an inclined plane in space with the (x,y,z) coordinates of each vertex being those of a sample point. Like the Thiessen polygons, the sizes and orientations of the individual figures are functions of the spatial distribution of the sample points. However, with the TIN, each triangle models the change from sample unit to the next in terms of an assumed linear transition between the z values.

To create the TIN from a set of sample points requires an algorithm to decide which neighboring points get connected to form the triangles. Of the many ways to approach the problem, the Delaunay triangulation has been shown to have the desired characteristic of a unique solution for any orientation of a set of sample points. Furthermore, this algorithm assures that the triangles formed are the "most equilateral" possible, and that a circle which circumscribes a triangle will contain no other sample points (McCullagh, 1981).

A TIN is really a set of interconnected triangular planes. Viewed differently, it is a network of linear transitions from each sample point to its neighbors. Using simple linear computations, interpolation of intermediate points may therefore be performed directly from the triangles in which they are located. Alternatively, more elaborate interpolation algorithms include the characteristics of neighboring triangles as well. One such algorithm which was proposed by Akima (1978) seems to enjoy fairly widespread use. The interpolation involves the application of a fifth order (quintic) surface of 21 coefficients within each triangle of a TIN. Each local surface is fit by imposing 18 conditions to the z values and to the first and second derivatives at the vertices, and three conditions at the triangle's sides to maintain continuity where adjacent surfaces meet.

We should mention that there are numerous methods for analyzing spatial data to produce an estimated isopleth or contour map. We have chosen this methodology to focus on due to its previously mentioned effectiveness at modelling the three-dimensionality of terrain, and because of its widespread use.

Spatial Extrapolation

If an inventory design purposefully includes sample points along the population border, or even beyond, then extrapolation beyond the convex hull is probably unnecessary. As this is unlikely in most forest inventories, some form of extrapolation is needed.

A simple, if not trivial, approach is to use the selected interpolation algorithm to make estimates beyond the convex hull outward to the population border. This seems to be an undesirable approach because the results are data and algorithm specific, and it may lead to erratic, unrealistic surface estimates along the population border. The most unsatisfactory results are likely to occur with trend surfaces and TINs.

An approach that we have found to be reasonably successful starts by generating a set of artificial sample points outside of the population. The spacing and positioning of these points are such that a new convex hull will fully include the population. Values for z are assigned to each of these points as a function of the neighboring sample units that are interior to the population border. The interpolation model is then applied to the extent of this new outer convex hull. Finally the actual border of the population is superimposed to clip or trim the estimated surface at the limits of the population.

We have conducted several trials using differing techniques for the optimum selection of the artificial sample points and for the assignment of z values. We have experienced satisfactory results that are improvements over the simple unrestricted extension of the interpolational model. Simplicity and applicability to one's particular software would dictate which technique is most applicable.

Comparative Results

We have conducted a set of empirical studies to try to gain insight into the magnitude of differences between traditional statistical estimates and those generated by spatial interpolation and extrapolation techniques. The techniques that we used were Thiessen polygons, trend surfaces and TIN based contour mapping. The data set was a portion of the 1983 USFS forest inventory for Vermont

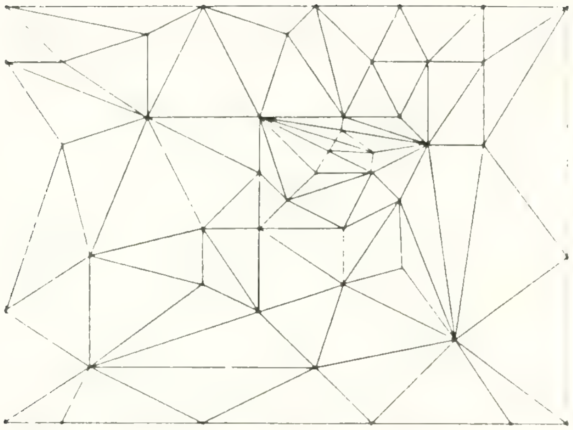


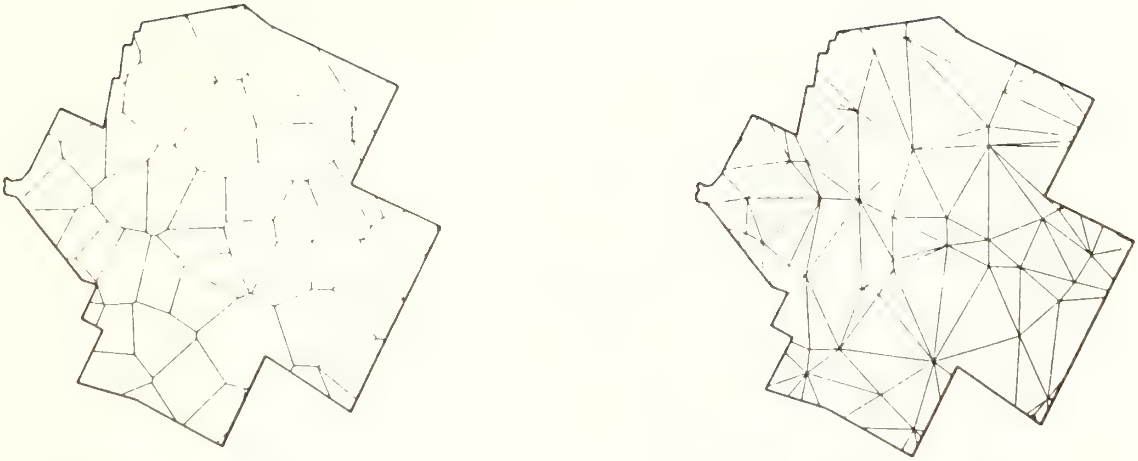
Figure 4. Triangular irregular network (TIN) (from Burrough 1986).

(Frieswjk and Malley, 1985). In order to consider differing spatial distributions and autocorrelations, we used several variables ranging from elevation to gross cubic foot volume per acre, to basal area per acre to number of trees per acre. Using an inverse square weight in the calculation of Moran's I, autocorrelations ranged from a high of 0.77 for elevation to a low of 0.07 for trees per acre.

We are just now completing the evaluation of the results for area estimates for arbitrarily selected categories. A visual comparison of the results from the different techniques for one county if Vermont is presented in figure 5. Initial evaluations suggest that across the range of variables trend surfaces have consistently yielded significantly different results from those for traditional statistics. The differences for Thiessen polygons do not appear to be significantly different, and those based on an evaluation of the TINs are mixed and apparently more variable or autocorrelation dependent. Evaluations of totals and means are yet to be completed. The results thus far, however, are consistent with our expectations. Individuals wishing more details of this study should address requests to the authors.

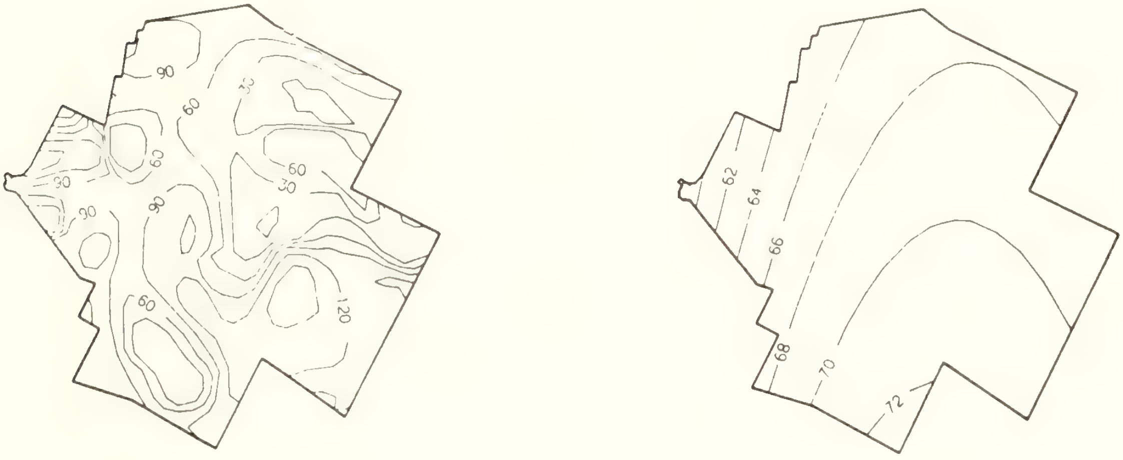
DISCUSSION

Our intent with this paper was to present a brief overview of some of the more common spatial analysis techniques that might be applied to forest inventory data to generate estimated surfaces and their related total and mean estimates. We believe that these types of estimates will be routinely done once the inventory is resident as a GIS data layer. It is important that, as inventory analysts, we understand the possible consequences of such applications. If statistical estimates and spatial analysis procedures are to yield compatible results, much work needs to be done in the development of error estimation techniques to assist in the evaluation process. Furthermore, there seem to be sufficient differences in estimates that the possibility of developing constrained analysis procedures should be investigated.



(a) Thiessen polygons

(b) Triangular irregular network (TIN)



(c) Contours generated from the TIN

(d) Contours for 3rd order trend surface

Figure 5. Visual compariosn of statewide spatial interpolation of bassal area per acre for Lamoille County, Vermont.

REFERENCES

Akima, H., 1978: A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points. ACM transactions on mathematical software, 4(2):148-159.

Amedeo, D. and Golledge, R.G., 1975: An introduction to scientific reasoning in geography. John Wiley & Sons, Inc. New York, 431 pp.

Averack, R and Goodchild, M.F., 1984: Methods and algorithms for boundary definition. In Proceedings of the international symposium on spatial data handling, Zurich Switzerland (D.F. Marble, K.E. Brassel, D.J. Peuquet and H. Kishimoto, editors), Geographisches Institut, Zurich, pp 238-250.

Burrough, P.A., 1986: Principles of geographical information systems for land resources assessment. Clarendon Press, Oxford, 194 pp.

Cliff, A.D. and Ord, J.K., 1981: Spatial processes: models and applications. Pion Ltd., London, 266 pp.

Cochran, W.G., 1977: Sampling techniques (3rd ed). John Wiley & Sons, New York, 428 pp.

Davis, J.C., 1976: Contouring algorithms. In Proceedings of the 2nd international symposium on computer-assisted cartography (R.T. Aangeenberg editor), American Congress on Surveying and Mapping, pp 352-359.

Frieswyk, T.S. and Malley, A.M., 1985: Forest statistics for Vermont -- 1973 and 1983. USDA Forest Service NE Sta. Resource Bulletin NE-87, Broomall, Penn., 102 pp.

Geary, R.C., 1954: The contiguity ratio and statistical mapping. The Incorporated Statistician 5:115-146.

Hsu, Mei-Ling and Robinson, A.H., 1970: The fidelity of isopleth maps. North Central Publishing Co., St. Paul, Minn., 92pp.

Johnston, R.J., 1980: Multivariate statistical analysis in geography -- a primer on the general linear model. Longman Group Ltd., London, 280 pp.

MacEachren, A.M. and Davidson, J.V., 1987: Sampling and isometric mapping of continuous geographic surfaces. The American Cartographer, 14(4):299-320.

Matern, B., 1947: Methods of estimating the accuracy of line and sample plot surveys. Medd. fr. Statens Skogsforsknings Institute, 36:1-138.

Matern, B., 1960: Spatial variation. Medd. fr. Statens Skogsforsknings Institute, 49:1-144.

McCullagh, M.J. 1981: Contouring of geological structures by triangulation. In Computer applications in Geology III, annual meeting of the Geological Information Group, London.

Moran, P.A.P., 1950: Notes on continuous stochastic phenomena. Biometrika 37:17-23.

Odland, J., 1988: Spatial autocorrelation. Sage Publications, Calif., 87 pp.

Peucker, T.K., Fowler, R.J., Little, J.J. and Mark, D.M., 1978: The triangulated irregular network. In Proceedings of the digital terrain symposium, Amer. Soc. of Photogrammetry, St. Louis, pp 24-31.

Pielou, E.C., 1977: Mathematical ecology (2nd ed.). John Wiley & Sons, New York, 385 pp.

Shepard, D.L., 1984: Computer mapping: the SYMAP interpolation algorithm. In Spatial statistics and models (G.L. Gaile and C.J. Willmott, editors), D. Reidel Publishing Co., Boston, pp 133-145.

Taylor, Peter J., 1977: Quantitative methods in geography. Houghton Mifflin Co., Boston, 386 pp.

Unwin, D., 1981: Introductory spatial analysis. Methuen & Co., Ltd., London, 212 pp.

Upton, G. and Fingleton, B., 1985: Spatial data analysis by example (vol 1). John Wiley & Sons, New York, 410 pp.

Spatial Analysis Using a Systematic Grid of Forest Inventory Plots: Selected Examples for the South Central United States

William H. McWilliams Research Forester
Joanne L. Faulkner Forester

USDA, Forest Service
Southern Forest Experiment Station
Box 906, Starkville, MS 39759

ABSTRACT

Spatial analysis facilitates the investigation of the location and extent of forest resources. Inventory plot information can be retrieved for any geographic area of interest within the South Central United States and can be characterized using either single or multiple attribute(s). Spatially displayed inventory plots are then linked directly to forest inventory statistics. Three examples are provided using simple spatial analyses and a systematic grid of forest inventory plots to illustrate the use of this technology.

INTRODUCTION

Shifts in product demand, complex resource interactions, and a diverse user constituency make forest resource analysis a formidable task in the Southern United States. Spatial analysis has emerged as a useful tool for addressing the many questions raised by the analyst. Coverage of the many applications of spatial analysis in forest inventory is beyond the scope of this paper. The intent is to illustrate some simple techniques for analyzing regional forest inventory data. Examples are provided using a systematic grid of forest inventory plots maintained by the USDA, Forest Service, Southern Forest Experiment Station, Forest Inventory and Analysis unit (SO-FIA), which conducts successive forest inventories in seven South Central states and Puerto Rico. The states include Alabama, Arkansas, Louisiana, Mississippi, Oklahoma, Tennessee, and Texas.

SAMPLING DESIGN

The SO-FIA uses a two-phase sample that consists of temporary aerial photo points and a grid of permanent ground plots. The area of forested land is determined by photointerpretation of temporary points and field checks of permanent plots. Field measurements are conducted on a subset of permanent plots. These measurement plots are located on a 3-by-3 mile grid, and so constitute a systematic or, as discussed by Ripley (1981) and Milne (1959), a "centric" systematic sample.

Tree data are collected on measurement plots that are classified as forest at the time of either the present or previous inventory. Each plot comprises a cluster of 10 satellite points. At each point, live trees at least 5.0 inches in diameter at breast height (d.b.h.) are tallied on a variable radius plot defined by a 37.5 factor prism. Saplings are measured on fixed plots centered around the first three points. Tree data are used to estimate volumes, biomass, basal area, numbers of trees, and relative stocking per acre. Plot level parameters, such as stand age and disturbance class, are also collected. Per acre estimates are expanded using factors derived as part of the forest area determination.

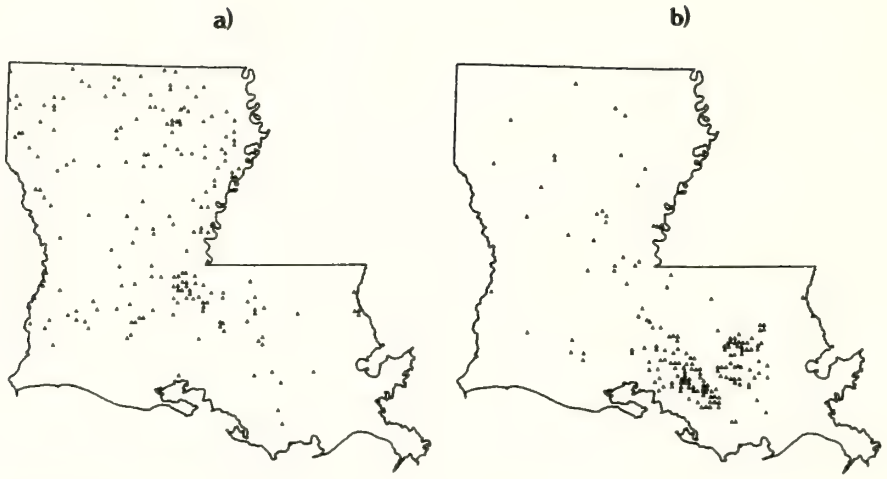


Figure 1. Distribution of a)sweetgum-Nuttall oak-willow oak and b)cypress-tupelo cover types in Louisiana. Each symbol represents an SO-FIA sample plot.

CHARACTERIZING SAMPLE PLOTS

Depicting regional forest inventory information requires large computer storage capacity and a flexible system for accessing spatial data. The SO-FIA sample for the current inventory cycle consists of 34,000 plots, roughly half of which are classified as forest. Both plot and tree level data are stored on the SO-FIA computer. Data access is facilitated by a relational data base management system (USDA Forest Service 1987). The analyst has several options for retrieving spatial data, including the entire region, individual states, groups of counties, and any polygonal or circular configuration of forest inventory plots.

Single Attributes

A common application of spatial analysis is to display inventory plots that have been classified according to single attributes. Single attributes can be derived using tree or plot information. One example of tree level classification is the species occurrence map. The map displays inventory plots where a particular species has been tallied. At the plot level, a map can be generated for all plots that occur in a given stand condition, such as sawtimber size plots.

To illustrate the single attribute approach, Figure 1 depicts the use of SO-FIA's cover type classification to analyze the vulnerability of bottomland hardwood habitat (McWilliams and Rosson 1989). The maps show the general distribution of sweetgum-Nuttall oak-willow oak and cypress-tupelo cover types in Louisiana. The two cover types account for roughly equal areas of forest land (1.2 million acres). Both cover types are threatened by landclearing activity that fluctuates with the prices of oil and agricultural products. The wider dispersion of sweetgum-Nuttall oak-willow oak forests makes them less vulnerable to endangerment than cypress-tupelo forests, which have a more clustered distribution. The impact on species diversity resulting from this potential loss can be assessed by compiling species composition information for the respective cover types.

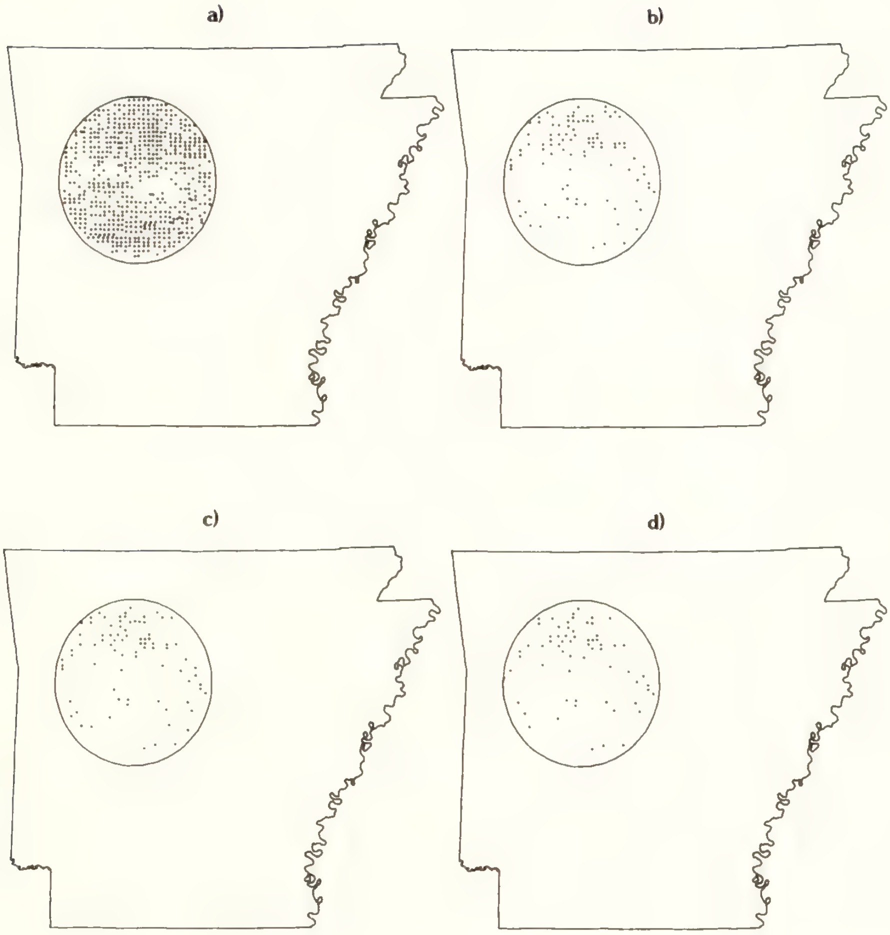


Figure 2. Distribution of SO-FIA sample plots within a 50 mile radius of Russellville, Arkansas, subject to the following successive constraints: a)classified as forest, b)contain at least 3,000 board feet (International 1/4 inch rule) of hardwood sawtimber per acre, c)within 0.5 miles of a road, and d)hardwood trees 15.0 inches and greater in diameter contribute at least half of the volume per acre. Each symbol represents an SO-FIA sample plot.

Multiple Attributes

Examining more than one attribute at a time highlights the interaction of variables. For example, the availability of hardwood sawtimber depends on several variables (McWilliams and Rosson 1988). Figure 2 demonstrates the impact of a set of four hypothetical constraints. The example is for SO-FIA inventory plots within a 50-mile radius of Russellville, Arkansas. Each constraint is evaluated simultaneously with previous constraints. The unconstrained set of plots represents 3.4 million acres of timberland that support 4.9 billion board feet (International 1/4 inch rule) of hardwood sawtimber. Imposing three availability constraints reduces the area represented to 0.4 million acres and 2.0 billion board feet. The spatial displays indicate the higher concentration of hardwood sawtimber in the northern part of the region and the relative impact of the availability constraints.

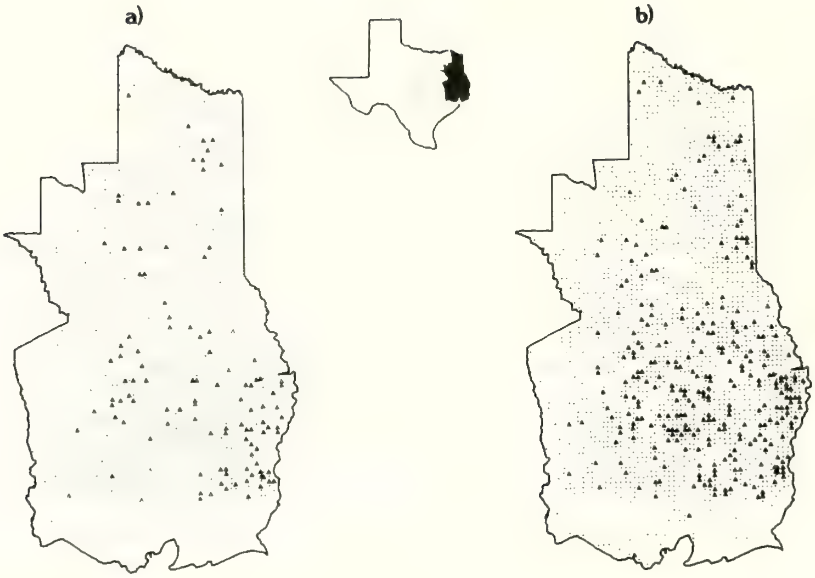


Figure 3. Distribution of pine forests in East Texas in a)1975 and b)1986. Each symbol represents an SO-FIA sample plot located in either a natural pine stand (dot) or a pine plantation (triangle).

Analyzing Change

Assessment of changes that occur between successive forest inventories is an important aspect of resource analysis. Very often, a change in one resource attribute relates to changes in several other attributes and can be used to draw conclusions about future trends. Figure 3 shows the extensive area of pine plantations that were established in East Texas between the two most recent inventories. The large scale conversion of older stands to young stands has helped explain decreases in softwood inventory and growth. Most of the plantations contain trees that are smaller than the 5.0 inch d.b.h. merchantability limit used by the SO-FIA. It is expected that the next forest inventory in East Texas will show significant increases in softwood inventory and growth since many of the plantations will have reached merchantable size.

SUMMARY

Spatial analysis of regional forest inventory data greatly expands the analyst's ability to answer questions pertaining to the extent and location of forest resources. It offers the capability to characterize inventory plots using single or multiple attributes and to analyze changes that occur over time. Spatially displayed inventory plots are linked directly to SO-FIA estimates of forest area, inventory, growth, removals, and mortality, thus providing a versatile means of quantifying forest resource information.

REFERENCES

McWilliams, W.H. and J.F. Rosson, Jr. 1988: Hardwood supply and availability in the midsouth highlands. In: Availability of Hardwoods in the Upland South, Hardwood Research Cooperative Series No. 7; 1988 January 7; Charlotte, NC. Asheville, NC: Hardwood Research Council: 47-95.

McWilliams, W.H. and J.F. Rosson, Jr. 1989: Composition and vulnerability of bottomland hardwood forests of the coastal plain province in the South Central United States. The International Forested Wetlands Resource: Identification and Inventory; 1988 September 19-22; Baton Rouge, LA. In: Forest Ecology and Management; Elsevier Science Publishers. (in press).

Milne, A. 1959: The centric systematic area-sample treated as a random sample. Biometrics. Vol. 15, No. 2, pp. 270-297.

Ripley, B.D. 1981: Spatial Statistics. New York: John Wiley and Sons. 252 p.

USDA, Forest Service. 1987: Interactive data access manual. Unpublished manual. Starkville, MS: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. On file with: Forest Inventory and Analysis, Box 906, Starkville, MS, 39759.

APPLICATION OF PERMUTATION TECHNIQUES IN ANALYZING SPATIAL PATTERNS

Robin M. Reich
Assistant Professor of Forest Biometrics
Colorado State University
Fort Collins, Colorado 80523

Paul W. Mielke, Jr.
Professor of Statistics
Colorado State University
Fort Collins, Colorado 80523

Frank G. Hawksworth
Supervisory Forest Pathologist
Rocky Mountain Forest and Range Experiment Station
USDA Forest Service
Fort Collins, Colorado 80526

ABSTRACT

Distance based multi-response permutation techniques are presented and discussed in the context of analyzing the spatial pattern of ponderosa pine stands infected with dwarf mistletoe. Under certain conditions, multi-response permutation techniques may provide useful information on spatial association and departures from randomness. Such information would be valuable in studying and evaluating the epidemiology, spread and intensification of dwarf mistletoe and other forest pests.

INTRODUCTION

The recognition of non-random patterns that exist in biological populations has led to the development of numerous mathematical models capable of detecting random, regular or aggregated spatial patterns (Clark and Evans 1954). A population, such as trees in a forest, is said to be randomly distributed if the location of a given tree is independent of the positions of all other trees in the same population. Aggregation, one of the most common spatial patterns in biological populations, occurs when individuals tend to form clusters due to competition, mortality and method of regeneration. In a plantation-type forest, individual trees are rather regularly distributed over a given area. The most common approach in analyzing spatial patterns rely on quadrat counts or distance measurements.

The quadrat method was developed primarily to study vegetative characteristics other than spatial patterns (Mawson 1978). Quadrat sampling has been used extensively in forestry, range, wildlife and ecology to sample such vegetative characteristics as frequency, density, abundance and presence (Curtis and McIntosh 1950). This method is also used effectively in detecting non-random patterns, but the results may vary with the size of the quadrat (Pielou 1957).

Unlike quadrat sampling, distance sampling was developed primarily to study the spatial relationship that exist within biological populations (Mawson 1978, Pielou 1959, Clark and Evans 1954). This method depends partly, or wholly, on distances from randomly selected points to the nearest individual, or from randomly selected individual to their nearest neighbor. In a randomly distributed population, it is assumed that the observed distance distribution follows a Poisson distribution.

At best such methods may be used to test departures from the underlying assumption of a random spatial distribution of elements over a given area. As an alternative, this paper explores the merits of using distance based permutation techniques for analyzing spatial patterns in ponderosa pine (*Pinus ponderosa*) stands infected with dwarf mistletoe (*Areceuthobium vaginatum*).

MULTI-RESPONSE PERMUTATION PROCEDURES (MRPP)

Let $\Omega = \{\omega_1 ,\ldots, \omega_N\}$ denote a finite population of N trees, $x_I = (x_{1I} ,\ldots, x_{mI})$ designate the m-dimensional coordinates corresponding to tree ω_I $(I = 1,\ldots,N)$, and let $S_1 ,\ldots, S_{g+1}$ be an exhaustive partitioning of the population into $g + 1$ disjoint groups such as healthy (H), infected (I) and dead (D) trees. Finally if we define $\Delta_{I,J}$ to be the Euclidian distance between ω_I and ω_J the distance based MRPP statistic can be defined as follows (Mielke 1986):

$$\delta = \sum_{i=1}^{g} C_i \xi_i \ ,$$

where

$$\xi_i = \binom{n_i}{2}^{-1} \sum_{I<J} \Delta_{I,J}\, \psi_i(\omega_I)\, \psi_i(\omega_J)$$

is the average distance between all distinct pairs of trees in group S_i $(i = 1,\ldots,g)$, $n_i \geq 2$ is the number of trees in group S_i $(i = 1,\ldots,g)$, $K = \sum_{i=1}^{g} n_i$, $n_{g+1} = N - K \geq 0$ is the number of unclassified trees such as dead trees in the excess group S_{g+1}, $\sum_{I<J}$ is the sum over all I and J such that $1 \leq I < J \leq N$, $C_i > 0$ $(i = 1,\ldots,g)$, is the proportion of the population in group S_i such that $\sum_{i=1}^{g} C_i = 1$, and $\psi_i(\omega_I)$ is 1 if ω_I belongs to S_i and 0 otherwise. To test the null hypothesis of no difference in the spatial distribution of trees MRPP assigns an equal probability to each of the

$$M = N! \,/ \left(\prod_{i=1}^{g} n_i! \right)$$

distinct allocations of the N trees to the g+1 groups. The collection of all δ values associated with the M equally likely allocations forms the permutation distribution of δ under the null hypothesis. The P-value associated with an observed value of δ (say δ_0) is the probability under the null hypothesis given by $P(\delta < \delta_0)$. For small sample sizes $(M < 1{,}000{,}000)$ the exact P-value under the null hypothesis can be determined while for large sample sizes one can approximate the P-value using the exact moments of δ. Also under the null hypothesis, the distribution of δ is often negatively skewed. To compensate for this skewness, the distribution of the standardized test statistic $T = (\delta - \mu_\delta)/\sigma_\delta$ is approximated by the Pearson type III distribution (Berry and Mielke 1983).

SPATIAL ANALYSIS

The application of MRPP in detecting non-random spatial patterns requires the use of an external reference set with known spatial pattern $\zeta(\omega_N)$. This external reference set is similar to the theoretical distributions used in detecting non-random spatial patterns for quadrat and distance samples in that it provides a relevant reference distribution with which the observed spatial patterns may be compared. For example, if $\zeta(\omega_N)$ represents a population with a random spatial pattern, a small value associated with δ, the MRPP test statistic, would indicated that the observed spatial pattern is not due to random chance, but some unknown non-random force.

To illustrate this procedure, consider the six artificial populations depicted in Figure 1. Each population consists of 30 points covering an area of 100 square units. Populations 1a through 1e were produced using a clustering algorithm for generating populations with varying degree of aggregation (Reich 1980), while population 1f was generated using a Poisson process. The density of each cluster (r) is given in the figure. Next, a reference population with a random spatial pattern is generated using a Poisson process. The reference population is then combined with the observed population to form a population of $N = K = 60$ individuals with $g = 2$ groups and $n_1 = n_2 = 30$ individuals within each group. Table 1 summarizes the results of the analysis. If the significance level were set at $\alpha = 0.10$ we would reject the null hypothesis of a random spatial pattern for all populations except 1e and 1f, which are somewhat randomly distributed. As the degree of aggregation changed from a random spatial pattern (population 1f) to a highly aggregated one (population 1a) the P-value associated with the null hypothesis of a random spatial pattern decreased from 0.4798 to 0.2169×10^{-9}.

SPATIAL ASSOCIATION

In populations in which more than one species occurs, it may be of interest to determine how the spatial distribution of one species influences that of another. If a species coexists with other species, they probably have some effect on each other, whether it is favorable or unfavorable. Because of the possible interaction between species it may not be appropriate to analyze the spatial patterns on a species by species basis (Pielou 1977).

The classical method of evaluating the spatial association between two species is to apply a χ^2 test to a 2 x 2 contingency table where the cells of the table represent the frequency of observing one of four types of nearest neighbor relationships (HH, II, HI, IH). A small P-value associated with this test would suggest that the species have distinctly different spatial patterns from each other.

A similar type of analysis can be done using MRPP. Consider the four spatial patterns depicted in Figure 2. Each population consists of $g=2$ groups (i.e. healthy and infected trees) of $n_1 = n_2 = 19$ individuals. In 2a both species are randomly distributed, while in 2b the species are aggregated with individual species occurring in distinct clusters. The overall spatial pattern in 2c is identical to 2a except that the species are somewhat aggregated. In 2d the two species occur in three distinct clusters with an equal number of individuals from each species randomly distributed within clusters.

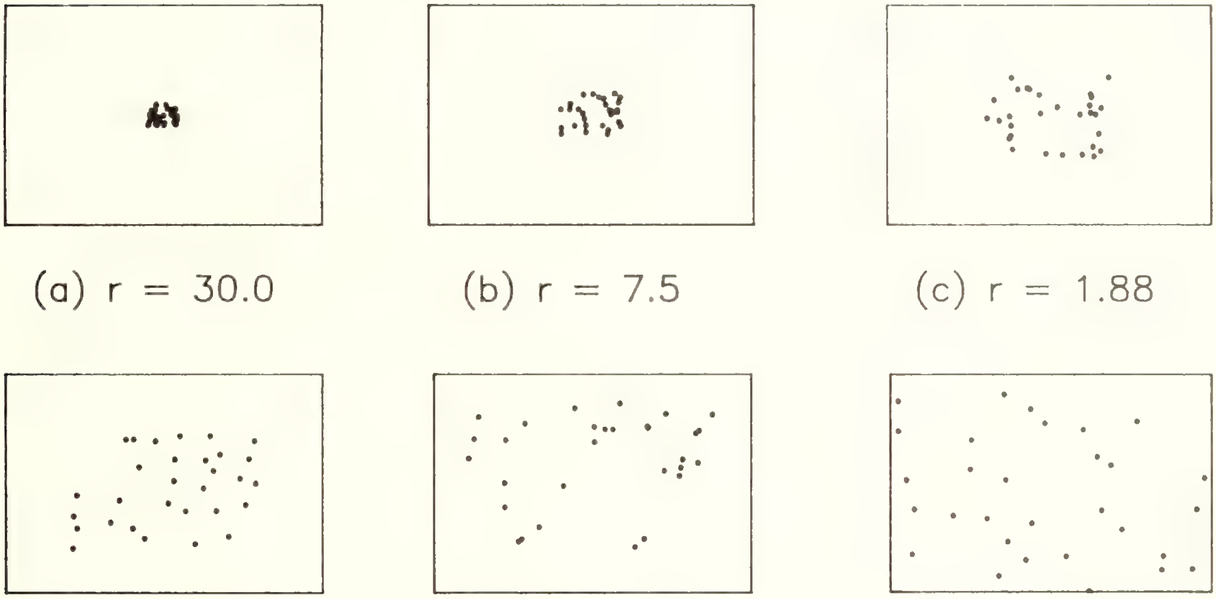
The results obtained for the four populations described above are summarized in Table 2. For population 2a the results suggest that the two species are arranged independently of one another (i.e. random spatial pattern) while in 2b the small P-value (0.6188×10^{-9}) associated with the test statistic indicates that the two species have distinctly different spatial patterns. In 2c the MRPP test statistic suggests that there is some segregation present between species ($p=0.1408$) but not as extreme as in 2b where there is no spatial relationship between species. In 2d the large P-value ($p=0.8784$) indicates that the two species have similar spatial patterns even though they tend to be highly aggregated.

In all of these examples, the MRPP test statistic is actually measuring the spatial association between species. In the case of 2a and 2d the MRPP test statistic suggests that both species have the same spatial relationship even though one population is aggregated and the other random. In the case of 2b the test statistic indicates that there is no spatial relationship among species, while in 2c the spatial relationship differed with respect the degree of aggregation and/or segregation.

In the case of 2c or 2d one may wish to take the analysis a step further and determine whether the observed spatial patterns differ significantly from a random spatial pattern. To test this hypothesis, population 2a which is assumed to have a random spatial pattern is treated as an excess group to provide an external reference set for comparing spatial patterns. When one compares the spatial pattern of 2c with 2a, the MRPP test statistic suggests that the two species are somewhat aggregated. If the same analysis is done for 2d the results indicate that the two species are randomly distributed, contrary

Table 1. Application of MRPP in determining the degree of non-randomness in a one-species population.

Population	Cluster Density	Test Statistic	Skewness	P-value
a	30.00	-17.061	-1.468	0.2169×10^{-9}
b	7.50	-12.619	-1.498	0.8969×10^{-7}
c	1.88	- 5.589	-1.670	0.8688×10^{-3}
d	0.83	- 1.544	-1.722	0.7914×10^{-1}
e	0.47	- 1.148	-1.767	0.1196×10^{-0}
f	0.30	0.239	-1.804	0.4798×10^{-0}



(a) r = 30.0

(b) r = 7.5

(c) r = 1.88

(d) r = 0.83

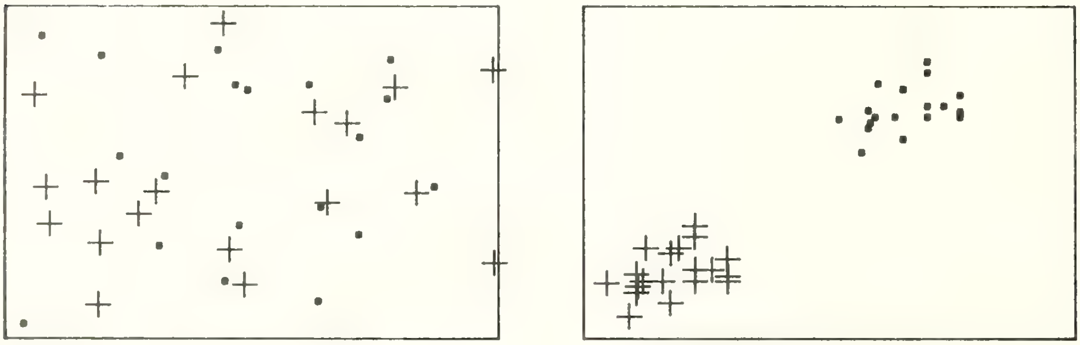
(e) r = 0.47

(f) r = 0.30

Figure 1. One species populations with varying degrees of aggregation, or cluster density (r).

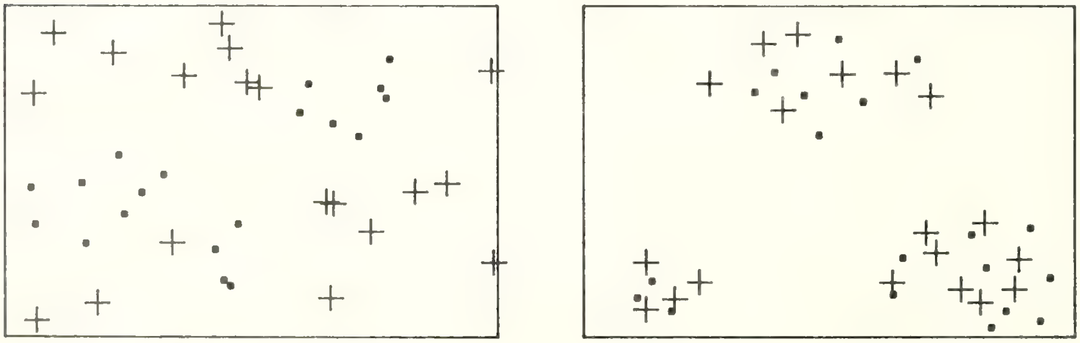
Table 2. Application of MRPP in analyzing the spatial relationship among two species.

Population	Test Statistic	Skewness	P-value
a	1.032	-0.916	0.8649
b	- 25.096	-2.659	0.6188×10^{-9}
c	- 0.992	-1.754	0.1408
d	0.895	-1.888	0.8784
c+a	- 1.699	-0.156	0.0489
d+a	0.676	-0.183	0.7453



(a) random

(b) aggregated



(c) random (aggregated)

(d) aggregated (random)

Figure 2. Two species populations with (a) random spatial patterns, (b) aggregated pattern in which the species are segregated, (c) both species together have a random spatial pattern although the species are aggregated, and (d) clumped pattern but the species are unsegregated.

to what we believe the true spatial pattern to be. This is due to the fact that the MRPP test statistic takes into consideration all distances between trees when evaluating spatial patterns, and not just distances between nearest neighbors. In the case of 2d the distribution of distances, between and within clusters, just happens to be similar to that of the random spatial pattern. This may be due to the fact that the clusters are approximately of equal size and equal distance apart, with the same number of species within clusters.

The multi-response permutation procedures presented here have some limitations. The first is that equal weight is given to all distinct pairs of distances between individuals within a given group. In practical problems where pine trees infected with dwarf mistletoe tend to infect their neighbors, it seems appropriate to vary the strength or weight of distances between trees based on their condition (healthy or infected) or size of individual trees. A second limitation is that an external reference distribution with a known spatial pattern must be defined in order to detect non-random spatial patterns. Since the MRPP statistic is based on distances between individuals within groups, changing the density of the reference population will influence the outcome. Care must therefore be used in defining the external reference distribution. This may not be a limitation but perhaps something that needs to be considered when using MRPP to detect non-random spatial patterns.

Although MRPP was originally designed to correct for some of the inherent problems (e.g. robustness) in some of the more classical statistical techniques, preliminary results indicate that when properly applied, this technique provides useful information on the spatial association and distribution of ponderosa pine trees infected with dwarf mistletoe. Such information should provide valuable insight in understanding the spread and intensification of this and other forest diseases.

REFERENCES

Berry, K.J. and P.W. Mielke. 1983: Computation of finite population parameters and approximate probability values for multi-response permutation procedures (MRPP). Commum. Statist. Simulation Comput., vol. 12, pp.83-107.

Clark, P.J. and F.C. Evans. 1954: Distance to nearest neighbor as a measure of spatial relationships in populations. J. Ecol., vol. 35, pp.445-53.

Curtis, J.T. and R.P. McIntosh. 1950: The interrelations of certain analytic and synthetic phytosociological characters. J. Ecol., vol. 31, pp.434-455.

Mawson, J.C. 1978: A Monte Carlo study of distance measures in sampling for spatial distribution in forest stands. For. Sci., vol. 14, pp.127-139.

Mielke, P.W. 1986: Non-metric statistical analysis: Some metric alternatives. J. Statist. Planning and Inference, vol. 13, pp.377-387.

Pielou, E.C. 1957: The effect of quadrat size on the estimation of the parameters of Neyman's and Thomas' distributions. J. Ecol., vol. 45, pp.31-41.

__________ 1959: The use of point-to-plant distances in the study of the pattern of plant populations. J. Ecol., vol. 47, pp.607-613.

__________ 1977: Mathematical ecology. John Wiley and Sons, Inc. New York. 385pp.

Reich, R.M. 1980: An evaluation of distance sampling and ratio estimation: A case study in slash pine plantations infected with fusiform rust. Master of Science Thesis, University of Florida, Gainesville. 140pp.

Forest Inventory and Geographic Information Systems

Moderator: Toby Atterbury

Lee Pierce Herrington, Geographic Information Systems: Where Do We Stand?

Sarah Osborne, Growth Models and Geographic Information Systems

David L. Verbyla, Building Predictive Models from Geographic Information Systems

Geographic Information Systems: Where do We Stand?

Lee Pierce Herrington[1]
State University of New York
College of Environmental Science and Forestry
Syracuse, New York 13210

ABSTRACT

The GIS world is exploding. New hardware and software is announced almost daily. Local, state and the federal agencies are acquiring hardware and software in the belief that it will help them carry out their missions more efficiently and precisely. Colleges and universities respond with more courses and the inclusion of GIS courses in required curricula. This rapid expansion of GIS capability and availability raises some very interesting questions regarding the accuracy needs of the digital databases representing the natural resources. What accuracy is needed in the digital data used in GIS query and modeling systems for resources management and planning? Will accuracy requirements be different for different operations. For example, will the accuracy needed for forest operations planning be the same as that needed for a GIS based inventory system? What is the effect of error on the decisions made as a result of using these systems? Modeling in particular, with its many operations applied to the data sets, is susceptible to large amounts of absolute error (the pixel is not what it is supposed to be). However, the decisions made based on these inaccurate modeling results seem to make sense nonetheless. More information is needed on the effect of accuracy on the decision making process.

INTRODUCTION

There has been, over the last few years, an explosion of Geographic Information Systems (GIS) software and a lowering of the cost of hardware needed to run the software. This has put GIS technology within economic reach of many organizations which only a few years ago could not even consider using the technology. The acquisition of the technology has created a demand for graduates with the knowledge needed to apply it to organizational problems. This has in turn resulted in the creation of university courses and programs of study in GIS . However, in many cases we do not know how to efficiently and economically use the tools which GIS provides.

The purchase of GIS technology is in many ways analogous to buying a car. There are Lamborghinis and there are Jeeps. Although both can get you from A to B they are quite different. There are analogs of the Lamborghini and the Jeep in the GIS world. Below are some of the characteristics of these vehicles

TABLE 1
Characteristics of vehicles

Characteristics	VEHICLE	
	Lamborghini	Jeep
Cost	Very high	Low
Performance	Very high	OK
Aesthetics	Excellent	OK
User friendliness	Poor	Good
Driving school	Yes	No
Maintenance costs	High	Low
Chauffeur needed	Probably	No, but would be nice

---

[1] Professor of Resources Information Management, Faculty of Forestry.

The selection of a vehicle to use is going to depend first on a) the job it is to perform, and b) the nature of the roads where it is to be used. From the GIS standpoint the matching of need to system is clear enough -- at first glance in any case. I think there is a tendency to assume that our needs are more sophisticated than they need to be. Roads are data. What good is a Lamborghini if all you have to run it on is woods roads? What good is a high precision system if the data is of low precision?

In what follows I will briefly describe where we are re Software and Hardware , the vehicles we have to chose from, and then turn to questions concerning the application of the technology to resources problems to illustrate how poor data and insensitive models could influence system purchase decisions. Do we need Lamborghinis or Jeeps?

SOFTWARE

There are a large number of GIS systems available (Parker 1988). They come in two basic flavors: The Raster or Grid Cell based systems exemplified by the Map Analysis Package (**MAP**) and its many derivatives and point-line-polygon (**PLP**) systems exemplified by ARC/INFO by ESRI (Burrough 1986).

Raster Systems.

These systems use a simple data structure -- a group of cells or pixels make up a layer or coverage. Each cell can have one attribute. Layers of themes can be combined in a large number of ways - added, subtracted, cookie cut, etc., including some very powerful and sophisticated operations. Raster based systems are excellent for many spatial modeling applications, are easy to learn, are generally user friendly, and cost 100s of dollars. Their main failings are that they do not draw very nice maps, cannot carry out network analyses, and are not very good at answering queries about what is where. The inability to make acceptable maps can be a major failing if you need to present maps to the general public. In addition, they can consume at lot of computer storage since every pixel must be stored (see Burrough 1986.) These systems are the Jeeps of the GIS world. (See Table I).

Point-Line-Polygon Systems (PLP).

These systems are also called arc-node or vector systems (Burrough 1986). Point- Line-Polygon is a very descriptive terminology since in these systems all maps, as with most paper map products, are made up of points, lines, and/or polygons. These systems can draw maps that look like what maps are supposed to look like. They are very good for query uses and some spatial modeling purposes. One reason they are good at query problems is that they can generally store multiple attributes per object (P,L, or P) through a connection between the graphic object and a database, usually a relational database. PLPs can carry out network analyses on topologically linked lines. Data structures are compact and storage requirements can be modest. However, processing is much more complex, data entry and editing is more complex, and spatial modeling within a polygon is not possible. These systems cost 1000s of dollars and generally require workstation quality graphics.

Other systems.

The GIS world is not quite as simple as described above. In fact many systems now embody both point-line-polygon and raster technology and many systems have the ability to transfer data between the two types. Generally going from PLP to raster is easier than going from a raster to a PLP representation. This exchange ability is important. We have found that overlaying raster maps with lines representing roads, compartment boundaries, water courses, etc. greatly improve the public's acceptability of raster output -- the jaggies[1] fade into the background when overlaid with lines.

A raster based system using a special form of raster called Quad-Tree, which stores cells of varying sizes, can cut storage requirements and can form an excellent base for query systems (Burrough 1986). A recent entry into the field which holds some special interest for resources management is the Object Oriented system (based on OOPS). Since resources management's objects (forests, compartments, stands) are hierarchical in nature the hierarchical nature of OOS seem to be a natural ally.

---

[1] Jaggies: The stair-step appearance of boundaries and lines in raster systems

HARDWARE

In a way I suppose that we can thank the hardware manufacturers for the sudden increase in interest in GIS at all levels of government. Some people say that the introduction of the MACII at the low end of the workstation market brought the prices of the established workstations down. I doubt the decrease in prices was that simple, but in any case, the prices of powerful systems are down. The performance of the microcomputer was and still is increasing rapidly while the prices of workstations are still falling. The shake out of the industry has not really happened yet so hardware is plentiful and powerful, and even a relatively simple AT or MAC can do useful work in GIS.

The fact that the falling hardware prices seemed to trigger many GIS purchases illustrates the fact that many people don't understand that the hardware and software costs are just the tip of the iceberg. Many people quote 20% as the fraction of total cost connected with hardware and software in bringing a real, useful GIS on line -- and this total does not include data maintenance. I believe that there are many systems out there now that will not be used at anywhere near their potential because of the misunderstanding regarding data for these systems. If computers in general are garbage pipes (GIGO) GIS systems are garbage amplifiers.

APPLICATIONS

The following are some observations concerning the application of GIS technology to real problems. They relate to whether to buy a Lamborghini or a Jeep. The list is not complete by any means.

Data Quality.

In a recent research project I needed to put into a GIS the compartment and stand boundaries of a rather large forest holding being managed for multiple use. The objective of the project was to provide a planning decision support system for forest management. Each compartment and each stand within the compartment had its own, separate map. We needed a composite map of the forest containing all compartments and stands. Even the compartment maps, which were based on USGS topo quads, did not fit together well. There were compartment overlaps and areas not included in any compartment no matter how we tried to fit the puzzle together. Some errors were hundreds of acres in size. The stand maps did not fit together very well either. Creating the boundary coverage was like putting together a jigsaw puzzle where several puzzles sawed out by several people were all mixed together. When we put these maps together as best we could there were numerous overlaps and underlaps (an overlap occurs when two or more stands occupy the same space and an underlap occurs when an area of the composite map is not included in any stand). Some of the errors, 200 acres in size, amounted to about 10% of the area of the stands. In addition, there was NO information on how, when, or by whom the stand maps had been created. These maps had been used for years without anyone, as far as we know, questioning their accuracy. This is not surprising. The maps were adequate for the function they were designed to support (see Hemenway 1989).

Two solutions were possible. The entire forest could have been resurveyed and thereby much more accurate boundaries could have been obtained, or we could live with the error and use a little poetic license and whatever information we had from the topo sheets and other sources to make compartments and stands fit. For our purposes this second option was fine. We had no need to know the exact area nor the exact location of the stands. We did need to know where they were relative to the general topography and did require that they be in the correct topological relationship to one another. The spatial modeling we planned to do was not, we thought, going to be sensitive to errors in boundary location. This view of data accuracy need is not universal, however, and some authors take the opposite view ((MacDougall 1975), for example). Part of the problem here is that we do not really know how data accuracy effects the drawing of conclusions and other resulting decisions made.

I believe that much of the base data we are now using in forest management is going to be like the data described above. The forest management profession had no need for highly accurate maps. In fact they were probably not aware of the inaccuracies in the maps -- it just was not important. When you incorporate this data into a GIS, however, the errors become obvious and unaesthetic. An expensive fix is to insist on highly accurate maps. Our solution was to fudge the maps.

The problem with fugding is that later on someone is going to used these maps in a land sale or some other operation which *is* sensitive to boundary error. If we did the job correctly and created what I call a pedigree for the map which would always be attached to the map this kind of misuse could be minimized (Grady 1988).

In this case, which I think is going to repeated again and again as forest management starts using GIS more and more, the data does not justify the use of the Lamborghini! Since the spatial data is known to be inaccurate why produce map products which imply a high degree of accuracy? The pixely maps of a raster based system would better reflect the nature of the data.

Spatial Modeling.

What is going to be done with the spatial and other data once it is in the GIS? The projects I have been associated with over the last several years indicate that the results of applying spatial modeling procedures to inaccurate data still yields correct decisions. The effect of uncorrelated error in coverages is significant. If each of several maps which are going to be overlaid and/or combined in a number of different ways has 10 or 20% random error then the result could easily contain 50% or more cells with incorrect attributes. The interesting thing is that, for many DSS problems for forest management I don't think that kind of error is necessarily going to lead to invalid conclusions and decisions. The problem is I don't ***know*** how much error of what kind measured by what standard will effect the decisions made

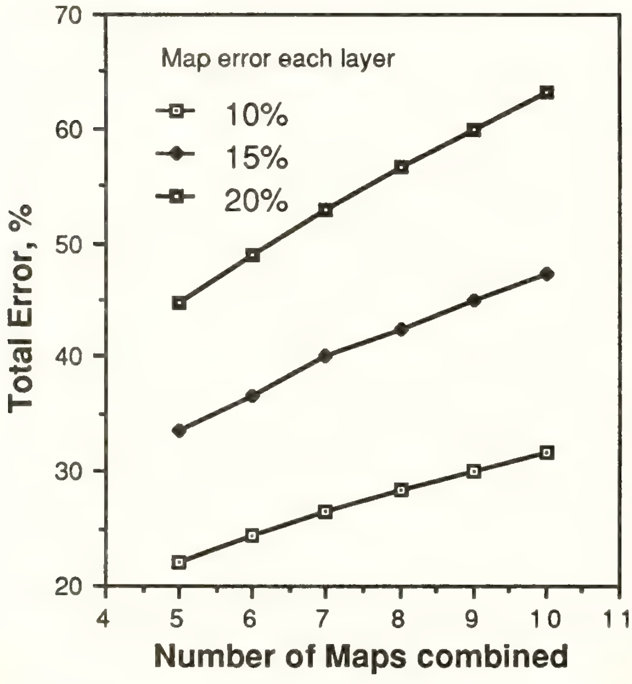


Figure 1. Accumulation of error for layers with different random error by number of maps added.

Susannah McKnight studied the effects of forest management practices in a mountainous area of Japan on a protected goat population (McKnight 1988). She needed maps of landuse and goat population at two times 10 years apart . The goat population data was of questionable accuracy because the method of collection varied from place to place. She had maps of landuse from one period which formed the base for surrogate forest management maps. The maps for the second period were based on forest management maps which formed the basis for surrogate landuse maps. This was the only data available. After processing this and other data through a rather complex spatial model she arrived at a set of maps that showed a definite interaction between goat migration and forest practice. This relationship could be seen clearly when you stood back from the map. Close inspection showed a lot of questionable cells. There was obviously a lot of noise in the final maps. Looking closely, however, was a classic case of not seeing the forest for the trees! If nothing else, this application of GIS technology using very shaky data could form the basis for a series of short, inexpensive studies to confirm or deny the GIS based conclusion.

The second case is described in more detail in the paper by Obsorne in the conference (Osborne 1989, Ferlow 1984, Herrington and Koten 1988). In this study a spatial model is used to help make decisions concerning which compartment should be harvested a any given time. I am convinced that boundary errors of several hundreds of meters, as long as they do not change any of the topology of the forest, would have little effect *as long as the area of the compartments was known.* Again, however, I don't have concrete evidence that this is so. We plan to run this model with known amounts of error and see at what point the decisions made from the analysis would differ from those made with the "correct" original data set.

If it turns out that many of the problems that we face in forest management are not severely effected by locational or attribute error then there is little need to use expensive, hard to use, systems. The Jeeps will do!

The worrisome part of this conclusion is that maps produced by the inexpensive Jeeps, which tend to be raster based systems, are very hard to use when the results must be presented to the public. Not only do unknowledgeable people not believe things which don't look like traditional maps, particularly if their house in not in the right place, but opposition lawyers would have a field day with the inaccuracies. It may be that we will be forced into using the Lamborghini even if it is not really needed and even if the Jeep would do just as well from a management point of view.

SUMMARY

Table 2, a conversion of Table 1 to a list of GIS characteristics for **PLP** and raster type systems, shows clearly the basic differences in the two systems. However, this is just one of many aspects of GIS that must be considered when specifiying a GIS for an organization. There are a number of papers dealing more fully with the problem of GIS selection (Reed 1988, Parker 1988, Jackson 1988, Burrough 1986 for example). There is a real need to develop knowledge about the effect of errors in GIS coverages on the conclusions drawn and decisions made as a result of manipulating these layers.

TABLE 2
Characteristics of the two basic types of GIS

Characteristics	Type of GIS	
	Point-Line-Polygon	Raster
Cost	Very high	Low
Spatial analysis Performance	Poor	Very High
Query Performance	Very high	Poor
Modeling capability	Moderate	Very Good
Accuracy	Can be very good	Poor
Aesthetics (acceptable maps)	Excellent	Poor
User friendliness	Poor	Good
Extensive Training needed	Yes	No
Maintenance costs	High	Low
Operator needed	Probably	No, but would be nice

CONCLUSIONS

Much of the spatial data that exists concerning our forest holdings, public and private, is inaccurate because we never needed a high degree of accuracy for forest management decision making -- and probably still don't need it from a strictly operational point of view. I think than many of the spatial models which will be developed to support forest management will not be sensitive to many kinds of errors and that decisions made from "noisy" maps will not differ from decisions made from highly precise, accurate maps. We need to investigate these questions before we go out and buy systems. We don't want to buy a Lamborghini when a Jeep will do.

REFERENCES

Burrough, P. A. (1986). Principles of Geographic Information Systems for Land Resources Assesment.Monographs on Soil and Resources Survey. Oxford, Clarendon Press.

Ferlow, N. (1984). Determining the Maximum Potential Stumpage Value Using a Digital Geographic Database. MS Thesis, College of Environmental Science and Forestry, Syracuse.

Grady, R. K. (1988). Data Lineage in Land and Geographic Information Systems. GIS/LIS '88. **2:** 722-729.

Hemenway, D. (1989). GIS Observer. Photogrammetric Engineering and Remote Sensing. **54:** 1505.

Herrington, L. P. and D. E. Koten. (1988). A GIS-Based Decision Support System for Forest Management. GIS/LIS'88. **2:** 825-831.

Jackson, J. (1988). Implementing a Low-Cost GIS for Transportation Planning. GIS/LIS'88. **2**: 932-940.

MacDougall, B. E. (1975). The Accuracy of Map Overlays. Landscape Planning. GIS/LIS'88.**2**: 23-30.

McKnight, S. K. (1988). The Japanese Serow: Wildlife Conservation and Management Through Geographic Information systems. MA Thesis. U. Cal. Los Angeles.

Osborne, S. (1989). Growth Models and Geographic Information Systems. State of the Art Methodology of Forest Inventory. **1:**

Parker, H. D. (1988). A Comparison of Current GIS Software: The Results of a Survey. GIS/LIS'88. **2:** 911-914.

Reed, C., N. (1988). A Minimum Set of Criteria for Selecting a Turn-key Geographic Information System: An Update. GIS/LIS'88. **2**: 867-872.

Growth Models and Geographic Information Systems

Sarah Osborne

Systems Design Specialist
GeoSource
6601 Kirkville Road
E. Syracuse, NY 13057

ABSTRACT

When forest managers develop a harvest schedule to maximize profitability, they are dealing with a situation that is both spatial and temporal in nature. Geographic information systems (GIS) are ideally suited to handle both dimensions. In this paper the feasibility of integrating a forest growth model with a GIS was tested using as a data base the SUNY College of Environmental Science and Forestry's Heiberg Forest. The product is a fully automated model that a forest manager can use as a management tool to determine the profitability of a forest stand at a point in the future as well as compare the state of the forest between several time periods. The basis for assessing harvest profitability is maximum potential stumpage which is the value of a stand at the mill after skidding and transportation costs have been subtracted. The final product is a fully automated model that a forest manager can use as a mangement tool to determine the profitability of a forest stand at a point in the future as well as compare the state of the forest between several time periods.

INTRODUCTION

Many industrial forest resource management decisions are both spatial and temporal in nature. In spatial terms, the manager may be looking for areas of a forest to harvest that are close to one another rather than spread throughout the forest. From a temporal standpoint, the manager wants a planned flow of timber for harvest. In this paper, the methodology of integrating a forest growth model with a geographic information system to create a model that has both a spatial and temporal dimension for assessing harvest profitability is discussed. The purpose of the model is to give a manager a visual tool that provides immediate information on harvest profitability at any point in the future as well as quantitative data to help in the analysis. The final product is a fully automated model that allows a forest manager to test different scenarios.

THE DATA SETS

The model requires two kinds of data sets. A spatial data set is needed which includes geographic information about the forest. A temporal data set is needed to describe how the forest changes over time.

The Spatial Data Base

The natural features of Heiberg Memorial Forest such as soils, hydrology, and topography were used in the spatial data base. Heiberg Forest, owned by SUNY College of Environmental Science and Forestry is located near Tully, New York, approximately twenty miles south of Syracuse, in the central part of the state (Figure 1). The forest sits on the northern escarpment of the Allegheny Plateau physiographic province, and the moderately rugged topography varies in elevation from 1300 to 2020 feet above sea level. The forest is made up of 97 compartments which include abandoned farmland, plantations of softwoods of various species, and second growth northern hardwoods.

Data for the Temporal Data Base

The temporal data base is created from FIBER 2.0, a stand projection growth model that predicts growth interactions among species in Spruce-fir, northern hardwood, and mixedwood forest types through five year time intervals (Solomon et al, 1986). FIBER is a two-stage matrix model. In the first stage the transition probabilities of tree growth and mortality as a function of stand density, tree size, and proportion

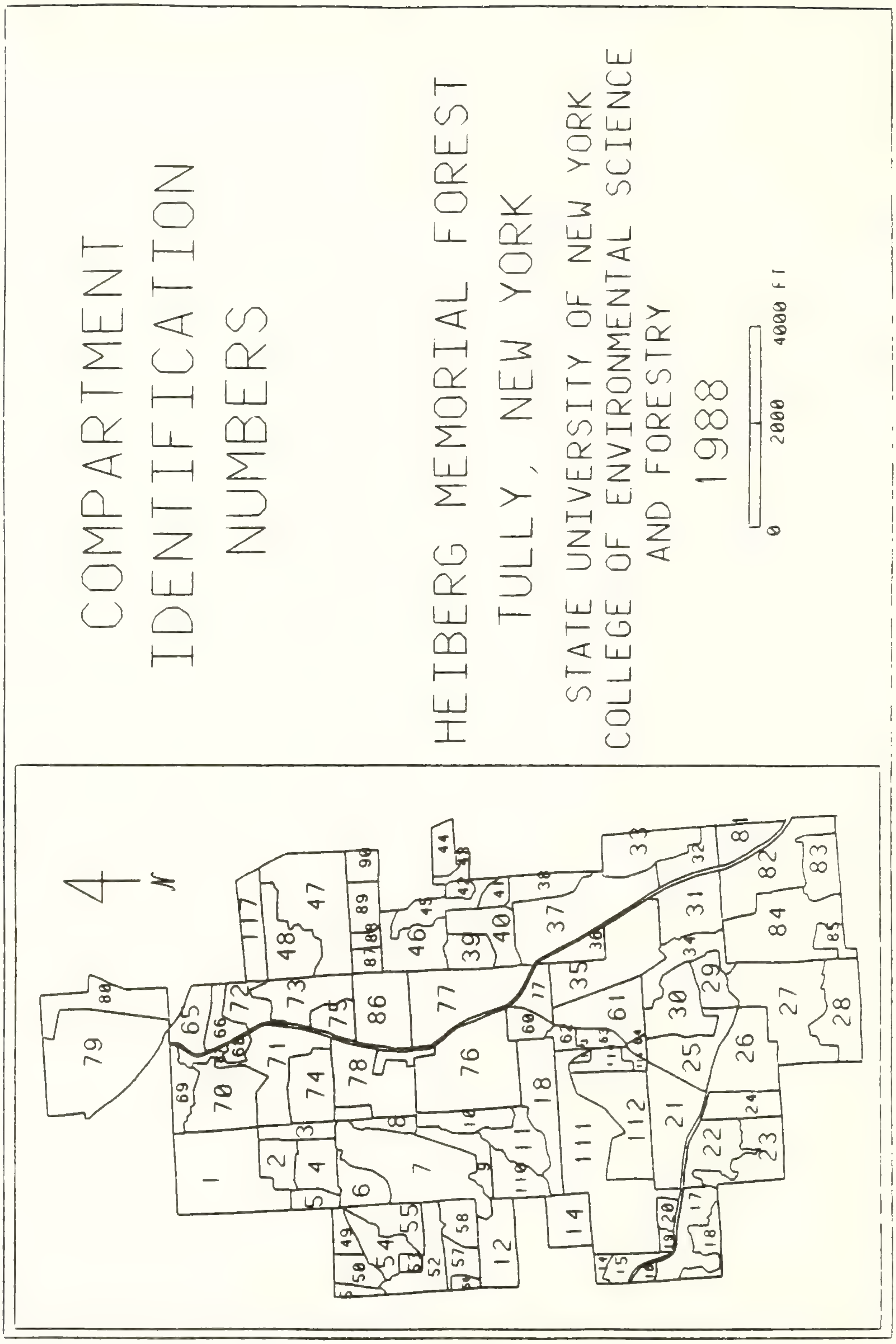


Figure 1. Map of Heiberg Forest by Compartment Number

of hardwoods is predicted through a set of linear equations. In the second stage the output matrices from the first stage are used to project the distribution of stand diameters. For Heiberg Forest, the FIBER model is applied separately to each compartment. The resulting stand volumes represent growth over time.

Compartment stand tables are the primary input to developing control files for the growth model. Since no tables are available for Heiberg Forest, representative "stand tables" were created based on data in the Heiberg Forest Continuous Forest Inventory (CFI). Additional information needed in the FIBER control files regarding stand management practices were taken from the Heiberg Forest Compartment Summary (June, 1987).

THE MODEL

The forest growth model/GIS integration builds on work done by Ferlow (1984) who calculated the maximum potential stumpage value (net stumpage value after skidding and transportation costs are subtracted) as a measure of harvest profitability. Ferlow used the Map Analysis Package or MAP (Tomlin, 1980), an inexpensive raster GIS with strong analytical capabilities, to generate a map that spatially represents maximum potential stumpage (MPS) values for Heiberg Forest. Input resource maps included forest type, soils, topography, roads, hydrology, and the compartment map. Additional non-spatial data included the cost to operate a skidder on each compartment and the volume of each forest stand. Assumptions behind the MPS values included the following: all merchantable trees within a compartment were harvested, skidding was always downhill to the nearest road, water could not be crossed, sawlogs were hauled to a hypothetical sawmill in Cortland, NY, and pulpwood was hauled to a hypothetical pulpmill in Syracuse, NY.

Three principal map layers were created from data in the original input maps: Market Value (MV), Cost-to-the-Landing (CTL), and Cost-to-the Mill (CTM). The final output map Maximum Potential Stumpage (MPS) was created by overlaying the map layers on a pixel-by-pixel basis in MAP using the formula:

$$MPS = MV - CTL - CTM$$

The Market Value map represents the mill-delivered price for wood. The value of each cell represents the stumpage volume per acre multiplied by the mill value of the forest type. The second map, Cost-to-the-Mill, was calculated by multiplying the stumpage volume by the cost of hauling sawlogs and pulpwood from the landing to the mill. The third map, Cost-to-the-Landing,represents the cost to skid logs to the nearest downhill road. Figure 2 is a summary of the inputs and outputs of the above maps (Herrington & Koten).

Several secondary maps were created in order to calculate Cost-to-the Landing values. One of the maps, Distance Cost, is the link between Ferlow's project and the growth model/GIS integration discussed in this paper. Distance Cost is a cost surface map that refers to the difficulty of skidding to the nearest road. Figure 3 is a flowchart showing inputs and outputs in obtaining Distance Cost.

Soils, topography, hydrology, and distance to the nearest road directly affect the ease with which logs can be skidded. A 'friction' map was developed which represents the relative difficulty of skidding. The higher the pixel value, the more difficult it is to skid within the area bounded by the pixel. The impact of soils, the difficulty of skidding on steep slopes, and the need for a skidder to bypass all water were reflected in a friction classification in which values may range from 0 to 200 (Ferlow, 1984).

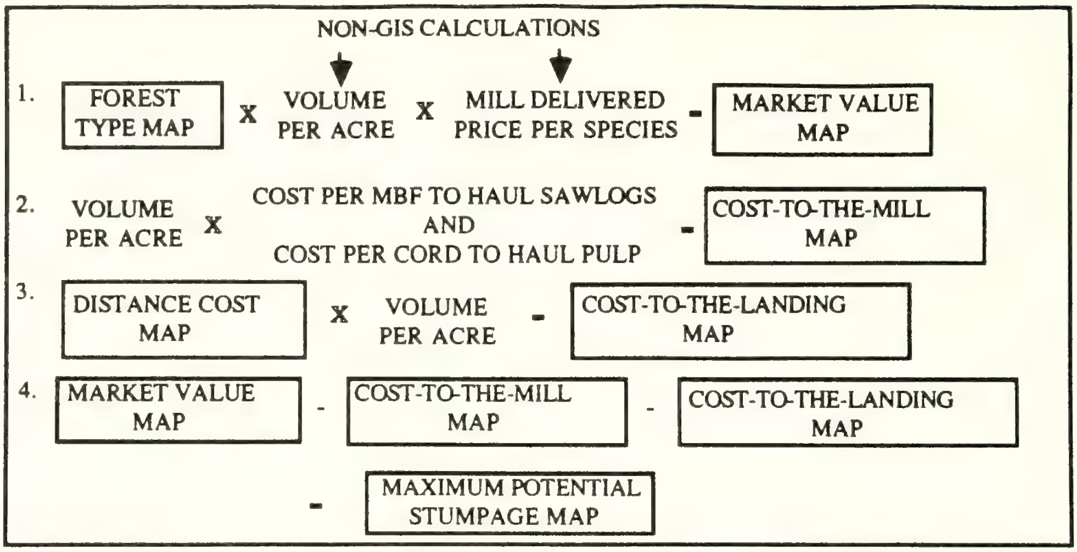


FIGURE 2. Summary of Inputs and Outputs to Create the Maximum Potential Stumpage Map

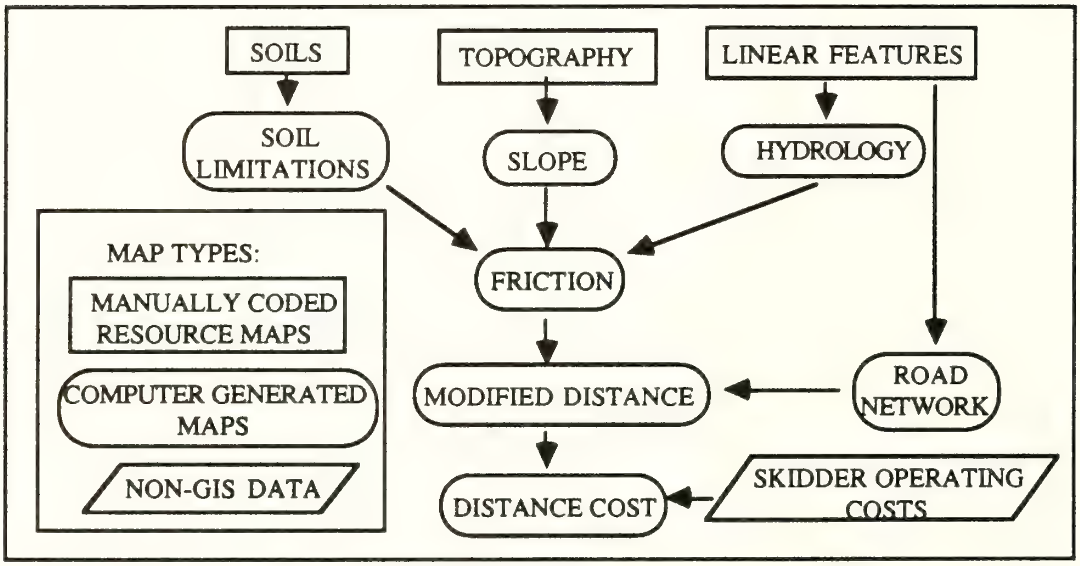


FIGURE 3. Flow Chart for Distance Cost

Using the friction and road network maps as inputs, another map, Modified Distance, was developed. The distance from each pixel to the nearest road was calculated using the path of least resistance downhill, ie. the path with the lowest summed friction values. The skidding path to the nearest road would be diverted around higher value pixels, and the distance value would be larger. The final map, the Distance Cost surface, is the modified distance pixel value multiplied by the cost to operate a skidder from that pixel to the nearest road.

The spatial modeling method used above to predict the profitability of harvesting is a snapshot of what exists at a given time. By adding the temporal dimension in the form of a forest growth model, harvest profitability may be predicted at any point in the future, and the state of the forest between two time periods may be compared (Osborne, 1989).

In the growth model/GIS integration model two GIS software packages are used in addition to FIBER 2.0: Idrisi which is a raster GIS similar to MAP, and pcArc/Info which is a vector GIS. The tree growth model, Idrisi, and Arc/Info operations are linked together through a series of Pascal programs.

During a model run FIBER first calculates stand volumes at a future time for each compartment. The process of calculating market value, skidding costs, transportation costs, and maximum potential stumpage values is similar to Ferlow's method except that the new stand volumes are used. Another difference is that while Ferlow calculated MPS on a pixel-by-pixel basis (each pixel represents one-fourth of a hectare), in this model all calculations are done on a compartment-by-compartment basis. In addition to maximum potential stumpage values, several temporal comparisons are made. The projected MPS is compared to present MPS in two ways: by percent value increase (a measure of percent growth) and by internal rate of return.

After all the values have been calculated for each compartment, six raster Idrisi map images are created: Market Value, Skidding Cost, Transportation Cost, Maximum Potential Stumpage Value, Percent Value Increase, and Internal Rate of Return. Once in Idrisi, the maps are converted to Arc/Info coverages. Arc/Info is used for display as it has a better graphic presentation and query system than Idrisi. In Arc/Info's graphics display mode, many map combinations can be displayed interactively based on user queries. Querying allows the user to play with what-if scenerios. Comparisons of the forest can be made using different criteria while maps showing the queried information are displayed immediately by the computer.

DISCUSSION

Combining spacial and temporal elements into a model that provides managers with a forest management tool is becoming more and more realistic from a practical and cost-effective standpoint. When this project of predicting harvest profitability was started in 1983, calculations were done on a mainframe computer with unpolished GIS software. Desktop computers were available but not commonplace. Six years later personal computers are very accessible, GIS software is more sophisticated, and the cost of running a GIS model such as the one described in this paper continues to decrease.

Data acquisition will continue to be the major expense in setting up geographic information systems. A GIS requires reliable maps and reliable information for its data base. One of the biggest problems in this project was the lack of data to build the basic stand tables for Heiberg Forest. The stand table data is a critical component since it is the source from which FIBER calculates stand volumes. These sawlog and pulpwood volumes are used as an input to every component in the maximum potential stumpage equation (market value, transportation to the mill, skidding costs, and maximum potential stumpage).

The description of the forest growth model/GIS integration provides a basic framework for adding more complex and realistic growth models depending on the species composition of a forest. In fact, the model has been set up so that several growth models could be plugged in quite easily. Any growth model which creates stand tables similar to those produced by FIBER can be used. FIBER was a viable model to test whether or not a forest growth model can successfully be integrated into a geographic information system. However, Heiberg Forest is on the fringe of the geographic range for which the FIBER model was developed. In addition, some of the species in Heiberg Forest are not modelled by FIBER. In these cases a similar species was substituted. As a result, although some of output was unrealistic, a viable method for including temporal change in a forest management decision support system was created.

SUMMARY

The model discussed in this paper provides one approach to dealing with both the spatial and temporal elements of a geographic information system. A model was developed to test predicting harvest profitability of a forest at a point in the future as well as compare the state of the forest between two time periods. While the model works it could be improved by inputting better stand volume data and including different growth models for different kinds of stands and stand compositions. A user interface has been set up so that a forest manager, with a minimal knowledge of computers and GIS, can run the model and test what-if scenarios.

REFERENCES

Ferlow, N., 1984: "Determining the Maximum Potential Stumpage Value Using a Digital Geographic Database." Master of Science Project. State University of New York, College of Environmental Science and Forestry, Syracuse, New York.

Herrington, L.P. and Koten, D.E., 1988: "A GIS Based Decision Support System for Forest Management." GIS/LIS '88, Proceedings of the Third Annual International Conference, Exhibits and Workshops. American Society for Photogrammetry and Remote Sensing, Falls Church, Virginia, pp. 825-831.

Osborne, S., 1989: The Integration of a Forest Growth Model with a Geographic Information System. Master of Science Thesis. State University of New York, College of Environmental Science and Forestry, Syracuse, New York.

Solomon, D.S., Hosmer, R.A., and Hayslett, H.T. Jr., 1986: FIBER Handbook: A growth Model for Spruce-Fir and Northern Hardwood Forest Types. U.S. Government Printing Office, Washington, DC.

Tomlin, C.D., 1980: The Map Analysis Package. Yale School of Forestry and Envrionmental studies, New Haven, Connecticut.

Building Predictive Models from Geographic Information Systems

David L. Verbyla
Assistant Professor
Department of Forest Resources
University of New Hampshire
Durham, New Hampshire 03824-3589

ABSTRACT

There is a great potential in many natural resource applications for building predictive models from geographic informations systems. However, there are 2 basic problems that modelers should be aware of: prediction bias and mixed measurement scaled predictor variables. Models developed with many predictor variables relative to sample size have a high potential for prediction bias. Models with prediction bias appear to be accurate when used with the sample data but are less accurate when tested with new data. Prediction bias is illustrated by developing a classifier that correctly classified 35 of 37 sample cases. However, the classifier was developed with random numbers and was therefore misleading and useless. Thus, predictive models developed from geographic information systems should be validated before being accepted. The second problem of predictor variables measured at different scales (categorical, ordinal, continuous) is that many statistical models assume all predictor variables are measured at the same scale. Classification tree analysis is introduced as one promising statistical tool that can be used to develop models from geographic information systems that contain data layers with different measurement scales.

INTRODUCTION

There are many potential applications of predictive modeling within the geographic information system (GIS) environment. For example, a hardwood company might want a model that predicts prime black walnut plantation sites from digital soils data. A forest entomologist might want to model areas of potential high infestation from a introduced insect pest given digital elevation, vegetation type, and stand structure data. A wildlife biologist might want to map potential winter range of a particular species based on a model incorporating topographic, vegetation and escape cover data. A fisheries biologist might want to model survival class of stocked trout with lake physical and chemical data.

Thus the modeling problem can be generalized to developing a model that will predict classes that can be mapped from candidate predictor variables. The objective of this paper is address two problems with building predictive models within the GIS environment:

1. The potential problem of developing a model with prediction bias. A model with prediction bias will appear to be a accurate model but will predict less accurately if it is validated with new data.

2. The problem of mixed measurement scales of candidate predictor variables. Many statistical models assume one measurement scale, such as continuous or categorical. How do we develop a predictive model with a GIS that contains data layers of different measurement scales?

PREDICTION BIAS

Prediction bias (sometimes termed model error rate) can occur when a model is based on spurious sample relationships that do not exist within the population. The potential of prediction bias is great if there are many candidate predictor variables and the model sample size is small. For example, in linear regression, we can always build a model with a r^2 of 1.00 with 1-n predictor variables (where n is the sample size). Yet such a model would predict poorly when applied to new data. The problem of prediction bias has been discussed in many fields including soil science (McQuilkin 1976), forestry (Verbyla 1986), wildlife science (Rexstad et al. 1988), economics (Lovell 1983) and statistics (Efron 1983,1986).

I will now illustrate this prediction bias problem with a hypothetical example. Suppose we are interested in mapping pine site index based on 19 soil measurements that have been entered into a GIS. We randomly sampled 37 pine stand for site index class and then use linear discriminant analysis to develop a model to predict site class based on the soil measurements. Our model predicted 35 of the 37 sample stand's site class correctly (Table 1) and therefore we use the model to generate a map of pine site class for the entire forest.

Table 1. Predictive model developed from 37 hypothetical sample cases.

		Predicted site index class								
Site index class	Actual membership	30-39	40-49	50-59	60-69	70-79	80-89	90-99	100-109	110-119
30-39	1 stand	1	0	0	0	0	0	0	0	0
40-49	9 stands	0	7	1	0	0	0	1	0	0
50-59	3 stands	0	0	3	0	0	0	0	0	0
60-69	8 stands	0	0	0	8	0	0	0	0	0
70-79	3 stands	0	0	0	0	3	0	0	0	0
80-89	3 stands	0	0	0	0	0	3	0	0	0
90-99	6 stands	0	0	0	0	0	0	6	0	0
100-109	1 stand	0	0	0	0	0	0	0	1	0
110-119	3 stands	0	0	0	0	0	0	0	0	3

Overall classification accuracy = 94.6 percent
(35 of 37 sample stands correctly classified)

This modeling approach is wrong. The discriminant model was actually developed with sample cases that had predictor variables that were random integers (1-100) and randomly assigned site class (1-9). The model had high prediction bias because it was developed with many predictor variables and a small sample size. Increasing the sample size is often not economically feasible. In this example, for each additional sample stand we would have to sample 3 to 5 dominant or codominant trees per stand for site index determination. However, because of this potential for model prediction bias, all models should be validated before they are used to generate predictive maps from geographic information systems.

MIXED MEASUREMENT SCALES

Potentially important predictor variables within geographic information systems are often measured at different scales. For example, a natural resource agency might have a GIS with categorical data layers (vegetation type, bedrock class, topographic position), ordinal data layers (soil drainage, stream order, erosion hazard class), and continuous data layers (digital elevation, stand basal area, stand age). Some statistical models assume predictor variables from only one measurement scale. For example, linear discriminant analysis assumes that predictor variables are normally distributed, with equal group variances and covariances, and therefore should be continuous.

One promising technique for building predictive models with mixed measurement scale data is classification tree analysis (Breiman et al. 1984, Verbyla 1987). Classification trees are predictive models structured as dichotomous keys. Consider the hypothetical example in Figure 1. A forest manager wants to map all areas that are potential prime pine sites with a GIS. Many candidate predictor variables are available that might be important in predicting prime sites, however in this hypothetical example only soil depth and percent stone content are considered. In classification tree analysis, the value of each candidate predictor variable that best predicts class membership is determined by a computer program called CART (Classification and Regression Trees, Breiman et al. 1984.) In this case the values would be 34 for "percent stone content" and 39 for "soil depth". Next, CART selects the predictor variable that best splits the sample cases into the purest class membership. In this example, a split at "percent stone content" of 34 would result in 25 cases correctly classified. A split at "soil depth" of 39 would result in 22 cases correctly classified. Therefore the variable "percent stone content" is chosen for the first prediction split by the model. The process continues to choose predictor variables for subsequent prediction splits. In this example, the second prediction split is at a "soil depth" value of 39.

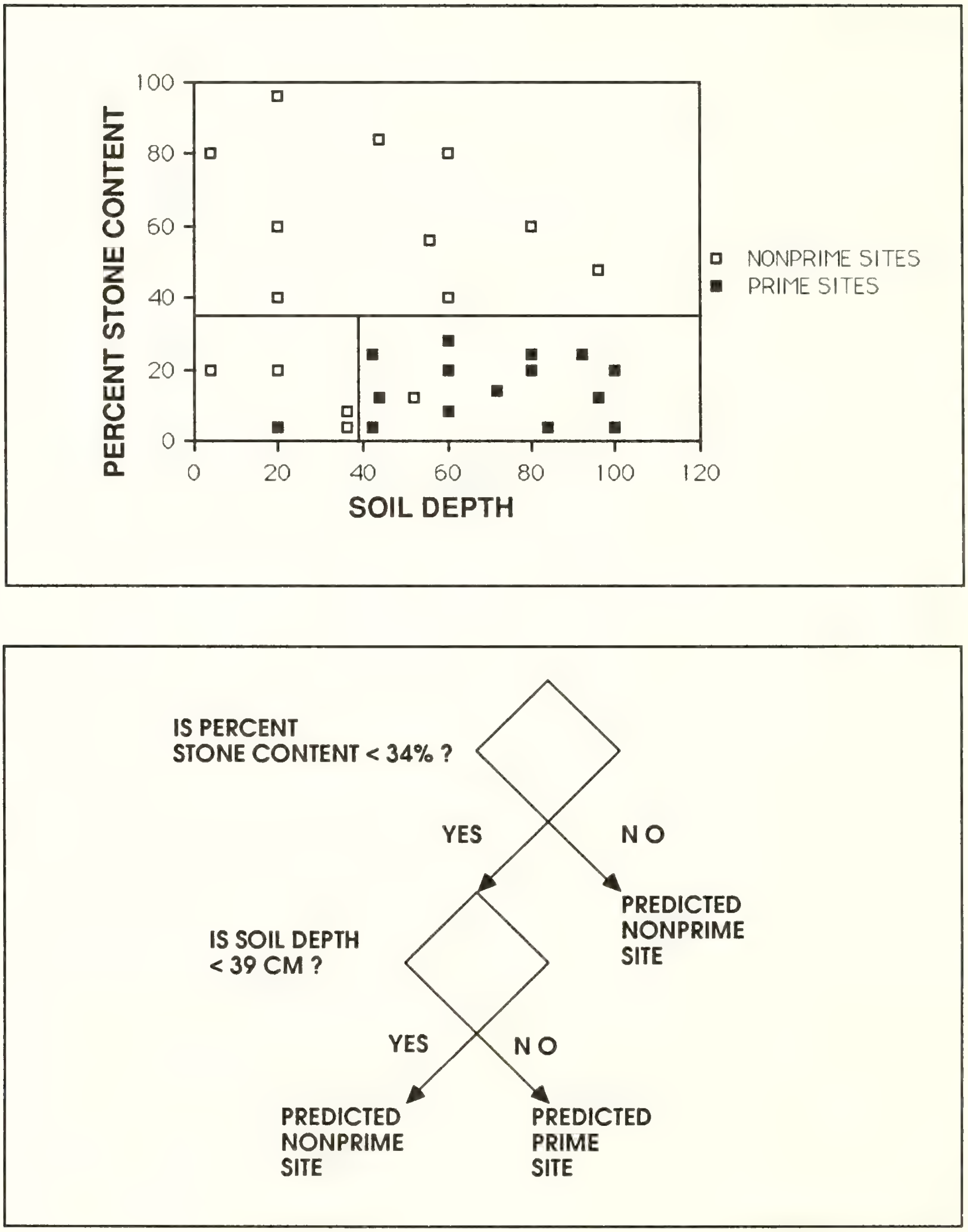


Figure 1. Classification tree developed with hypothetical data to predict prime sites.

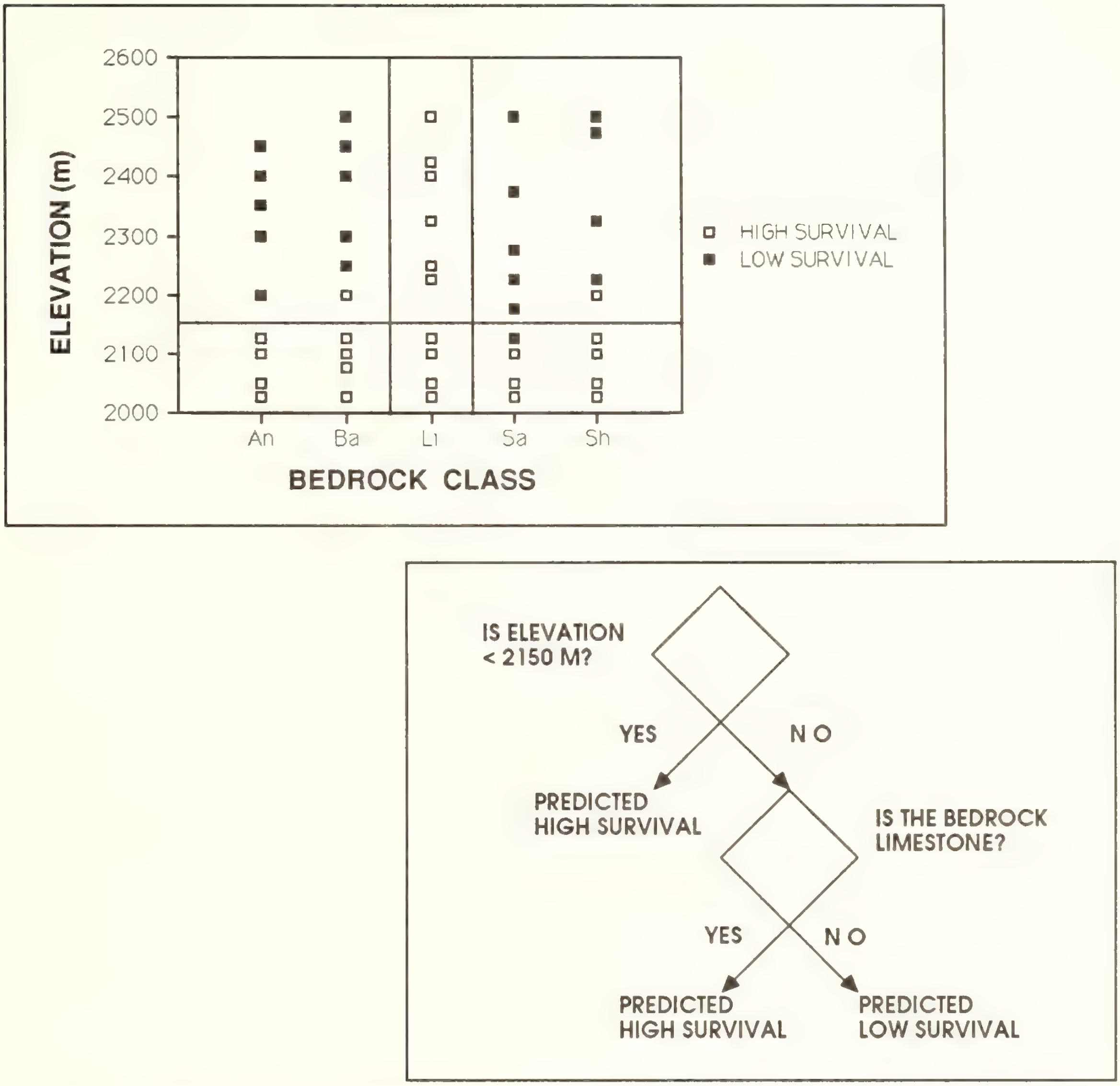


Figure 2. Classification tree developed with hypothetical mixed measurement scaled data (categorical and continuous predictor variables).

Classification trees analysis can easily be used with mixed measurement scaled data. Continuous and ordinal scaled data are analyzed by examining all splits of the form x(i) < C, where C is a constant ranging from the minimum value of x(1) to x(max). Categorical predictor variables are analyzed with an algorithm that examines all splits of the form x(i) where i is the ith subset. For example, we might have a GIS layer composed of 5 bedrock classes: An (andesite), Ba (basalt), Li (limestone), Sa (sandstone), and Sh (shale). The algorithm examines all 15 possible subset splits (x(i)=An, x(i)=Ba,. . . x(i)=Sa or Sh) and chooses the best possible split for the categorical predictor variable.

Figure 2 is a hypothetical example involving mixed measurement scales. A fisheries biologist wants a map of stocked trout survival class in an area of high acid deposition. The hypothetical classification tree predicts trout survival from a continuous variable (elevation) and a categorical variable (bedrock class). Keep in mind that in actual practice we often have many potentially important predictor variables and such a simple model is often not intuitively obvious.

Classification trees are easy to use with GIS overlay algorithms. They are also easy to understand because of their simple binary key structure. But in spite of their simplicity, they have a strong statistical foundation (Friedman 1977, Gordon and Olshen 1978,1985, Breiman et al. 1984) and can be as accurate as alternative statistical prediction models (Gilpin et al. 1983, Dillman and Koziol 1983, Verbyla 1989). They have been used sucessfully in medical research and have many potential GIS applications in natural resource management.

ACKNOWLEDGEMENTS

I thank Jim Barrett and Ernst Linder for reviewing this manuscript. This is scientific contribution 1632 of the New Hampshire Agricultural Experiment Station.

REFERENCES

Breiman, L., Friedman, J. H., Olshen, R. A. and C. J. Stone. 1984. Classification and regression trees. Wadsworth Publishing Co., Belmont, CA, 358 pp.

Dillman, R. O. and J. A. Koziol. 1983. Statistical approach to immunosuppression classification using lymphocyte surface markers and functional assays. Cancer Research, vol. 43, pp. 417-421.

Efron, B. 1983. Estimating the error rate of a prediction rule: improvement on cross-validation. Journal of the American Statistical Association, vol. 78, pp. 316-331.

_______. 1986. How biased is the apparent error rate of a prediction rule? Journal of the American Statistical Association, vol. 81, pp. 461-470.

Friedman, J. H. 1977. A recursive partitioning decision rule for nonparametric classification. IEEE Transactions on Computers, vol. C-26, pp. 404-408.

Gilpin, E., Olshen, R., Henning, H. and J. Ross Jr. 1983. Risk prediction after myocardial infarction. Cardiology, Vol. 70, pp 73-84.

Gordon, L. and R. A. Olshen. 1978. Asymptotically efficient solutions to the classification problem. Annals of Statistics, vol. 6, pp. 515-533.

_______________________. 1985. Tree-structured survival analysis. Cancer Treatment Reports, vol. 69, pp. 1065-1068.

Lovell, M. C. 1983. Data mining. The Review of Economics and Statistics. Vol. 65, No. 1, pp. 1-12.

McQuilken, R. A. 1976. The necessity of independent testing of soil-site equations. Soil Science Society of America Journal, vol. 24, pp. 783-785.

Rexstad, E. A., Miller, D. D., Flather, C. H., Anderson, E. M., Hupp, J. W. and D. R. Anderson. 1988. Questionable multivariate statistical inference in wildlife habitat and community studies. Journal of Wildlife Management, vol. 52, pp. 794-798.

Verbyla, D. L. 1986. Potential prediction bias in regression and discriminant analysis. Canadian Journal of Forest Research, vol. 16, pp. 1255-1257.

___________. 1987. Classification trees: a new discrimination tool. Canadian Journal of Forest Research. Vol. 17, pp. 1150-1152.

___________. 1989. An alternative approach to conventional soil-site regression modeling. Canadian Journal of Forest Research. Vol. 19, pp. 179-184.

Mathematical Models and Forest Inventory
Moderator: John C. Rennie

Harold E. Burkhart and Thomas M. Brooks, Status and Future of Growth and Yield Models

Alan R. Ek, Interfacing Forest Inventory With Growth and Yield Models

Edgar Kaufmann, Modeling With Data of the First Swiss National Forest Inventory

Status and Future of Growth and Yield Models

Harold E. Burkhart
Thomas M. Brooks Professor of Forest Biometrics
Department of Forestry
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

ABSTRACT

Growth and yield modeling methodology has advanced significantly in recent years, but the future promises even more rapid advancement. Progress in growth and yield modeling centers around the three key elements of data collection, analytical techniques, and computing technology. Presently, emphasis is placed on gathering data via permanent plots in operational stands, but more designed experiments are needed for the future. Data are being analyzed using sophisticated regression techniques, but the future promises a more "biomathematical" approach. Continued advancement in computing technology (both hardware and software) will allow enhanced output from growth and yield models and increased use of the models in computer systems for improved decision aids. Future modeling efforts will stress including an increased number of silvicultural alternatives, providing expanded information on tree quality and product yields, and allowing predictions of forest response to environmental stresses.

INTRODUCTION

Forecasting growth and yield is implicit in the process of forest management. Although yield tables for even-aged stands have been in use in the United States since about 1900, we still lack adequate growth and yield information for many purposes and circumstances. Consequently, growth and yield estimation continues to be an active area of research. While growth and yield forecasts enter into virtually all decisions, the primary uses of growth and yield information can be categorized as follows:

Inventory updating
Evaluation of silvicultural alternatives
Management planning
Harvest scheduling.

Growth and yield information is used for a variety of purposes; no single data base or modeling approach can be optimal for all purposes. As an example, if one were primarily interested in evaluation of silvicultural alternatives, designed-experiment type data with the relevant silvicultural treatments included would probably be used. The model structure would likely be quite detailed in terms of the structural equations and the types of output produced in order that the full range of treatments could be evaluated under varying assumptions. If, on the other hand, one were primarily interested in inventory updating, the data should be obtained from a representative sample of the types of stands to which the model is going to be applied. The input to the model would necessarily need to be consistent and compatible with the inventory data definitions and quantities available. The structural equations should be as simple and straightforward as possible for producing output (updated stand statistics) that is needed for and consistent with the inventory data base. What is "best" depends primarily on the objective(s) for developing the model; obviously, the objectives should be specified clearly before any data collection or analyses are initiated.

Questions arise with regard to whether we should be striving to develop multipurpose models that can be used for a variety of tasks or should we be developing a number of limited-purpose models. In choosing a growth and yield model one must be concerned with the stand detail needed for the particular decision at hand and the efficiency in providing this information. In situations where predictions are required for a very broad range of management decisions, it would be desirable to have a system of growth and yield models capable of providing logical and consistent estimates for varying degrees of stand detail (whole stand values, size class data, or individual tree information), thus allowing users to efficiently compute estimates with stand detail appropriate to the use of the information. A framework for such an integrated system of stand models was developed by Daniels and Burkhart (1988). However, many unresolved problems--such as what is the appropriate data base and parameter estimation method--remain, but work on integrating growth and yield models with varying levels of resolution is certain to continue because of the great consistencies and efficiencies that such integration can effect.

STATUS OF GROWTH AND YIELD MODELS

Growth and yield models produced to date can be characterized in the following manner:

1. Whole stand models
    a) aggregate values only
    b) size class information

2. Individual tree models
    a) distance independent
    b) distance dependent

The approach taken in each modeling effort has varied by objectives for developing the model, background and interests of the modeler, and the analytical methods, data base and computational machinery available. The three-way interaction of analytical methods-data base-computational technology has had a profound influence on growth and yield modeling. I will discuss the status of each of these three basic components and then, in the next section, will conjecture about future developments.

Data Collection

The typical approach taken in past growth and yield studies was to define a population of interest, obtain a sample from the defined population (the sample could consist of temporary plots, permanent plots or both), and estimate coefficients (invariably with least squares) in specified equation forms. This approach produces satisfactory prediction tools for many purposes, but it may not be adequate in circumstances where forest management practices and objectives are changing rapidly. Given that growth and yield models are used to project the present forest resource and to evaluate alternative treatment effects, data of both the inventory type (which describe operational stands of interest) and the experimental or research type (which describe response to treatment) are needed. The amount of effort that should be devoted to each type of data collection is not immediately obvious. Nor is it at all clear whether the data should be combined and a single model produced or the data kept separate and different models produced for different uses or objectives.

Analytical Methods

Nonlinear regression techniques are now widely used in growth and yield modeling. Easy-to-use computer software is readily available for nonlinear estimation. Thus we can expect use of nonlinear models and nonlinear estimation techniques to become even more pervasive in the future.

Growth and yield models often employ a system or set of equations to describe stand development. Generally the individual equations in the system are fitted one at a time by the method of least squares, but this may not be entirely satisfactory. The same variable may be dependent in one equation and independent in another. Further, certain coefficients in one equation may be functionally related to coefficients in another equation. And the residuals from the several equations are likely correlated. Because of these considerations, there is increasing emphasis in treating growth and yield models as a system of equations and simultaneously estimating all the coefficients in the systems. When simultaneously fitting equations to growth and yield data, problems can arise. Many of the systems in which we are interested involve nonlinear components and nonlinear constraints. While work has been done on simultaneous fitting of nonlinear models, considering nonlinear models generally makes the problem much more complex. Sample size may also be a problem. Although asymptotic statistical properties have been established for many simultaneous estimation procedures, the small sample properties are often unknown.

Computing Technology

Personal computers are playing an ever increasing role in the development and application of growth and yield models. In the recent past most of the work on development of growth and yield models was being done on mainframe computers. But data storage capabilities and statistical software availability has now reached a level to where development--as well as application--of growth and yield models is feasible on personal computers.

FUTURE

With that brief summary of the status of data collection, analytical methods, and computing technology as applied to growth and yield modeling, let's turn our attention to envisioning the future of these areas. While I have no crystal ball, nor do I make claim to unique insight into the future of these matters, I'll not let that stop me from speculating about future developments.

Data Collection

The most pressing data needs are for permanent plots in operational forest stands and for designed experiments. We have been limited by inadequate data in the past, and we continue to be data limited. Permanent plots established in the past have sometimes had limited usefulness because of inadequacies in the measurements taken. In any new permanent growth plots, I feel that the minimum data measurements should include dbh, height (on all trees), crown measures, stem quality assessment and tree spatial locations. More detailed data will be required to more fully exploit the enhanced computational facilities and analytical methods available to us.

Permanent plots alone will not solve all of the data needs. Carefully designed experiments are essential for testing hypotheses about stand dynamics and providing insight into model forms that should be fitted to large sets of permanent plot data representing forest populations of interest.

Permanent plots in operational stands, designed experiments at densities of interest to forest managers, and related types of data collection efforts require substantial investments of both time and money. While these types of data collection are extremely important, research is also needed on use of seedlings planted at very close spacing to speed up the process of stand development. Such "seedling stands" provide opportunities for minimizing site and environmental variability, reducing risks of loss of plots through catastrophic events, and producing initial data in a relatively short time.

Although relatively little work has been completed on using mini-rotation plantations for modeling forest stand growth and yield, there are past studies which show evidence of potential for future developments. An example is the model developed by Smith and Hann (1984) to describe the relationship between average plant size and stand density in single-aged, monospecific plant populations. The model was tested by growing red alder seedlings under greenhouse and lath house conditions at initial spacings of 8 x 8, 4 x 4 and 2 x 2 cm for 525 growth days. The model also adequately described the size-density trajectories for 20-50-year-old red pines growing at six initial spacings. If certain aspects of stand development are approximately equivalent between seedling and more mature stands, then we have a means for greatly shortening the time required to gain initial data. Caution is required in making inferences from seedling stands to mature stands. As stated by Ford (1984) "As a result of increasing size and biomass, a distinctive microclimate and soil environment develop which in turn strongly influence the growth process itself ...".

Analytical Methods

Future growth and yield analyses are likely to be more "biomathematical" in nature with less emphasis on empirical data fitting. As "biological laws" are developed and understood they are--and will be more so in the future--being explicitly incorporated into growth and yield models. Models that exhibit expected behavior with respect to biological understanding have greater generality and usefulness and inspire confidence on the part of users. Increased needs for using growth and yield models for making predictions for environmental and silvicultural conditions not included in the data bases emphasize the importance of aiming for more generality in model structure and approach.

Recent developments in so-called "experimental mathematics" --that is, fractal geometry and chaos theory--promise to influence future growth and yield analyses. While there is disagreement over whether chaos is merely an interesting idea enjoying a faddish vogue or a revolution in scientific thought (see, for example, the discussion by Pool 1989), some of the concepts and techniques will likely find application in efforts to quantify forest biology. Fractal geometry allows description of irregular objects such as tree crowns. Concepts from chaos theory that address such issues as sensitivity to initial conditions, stability, limits to predictability, should prove useful in the study of forest stand dynamics. Regardless of whether chaos turns out to be a revolution (i.e., a paradigm shift), as some claim, or a modest advance, as others believe, it deserves careful consideration by growth and yield scientists.

Computing Technology

Rapid advances in computing technology--both hardware and software--have had profound impacts on growth and yield modeling. Traditional uses of growth and yield information in decisionmaking will continue and new uses will expand. In order to make effective management decisions, biologic, economic, and geographic information must be integrated. Decisions are typically reached through economic analyses that are dependent on biologic and geographic characteristics of stands. Modern computing technology and geographic information processing software make the effective integration of biologic and geographic information feasible.

In the future, growth and yield models will be widely used in computerized decision aids such as expert systems. Concepts from the field of artificial intelligence will increasingly influence the user interface with growth and yield models.

Advances in computer graphics enable us to produce greatly enhanced output. The potential of computer graphics in analyzing forest dynamics is far greater than mere cosmetics, however. Graphic capabilities, coupled with increased computing power, open the doors to study of fundamental properties and behavior of systems of equations that were heretofore not possible.

Future Areas of Emphasis

After that brief, and I'm sure imperfect, glimpse of the future regarding data, analyses and computing, I will mention some areas that will surely receive increased emphasis by growth and yield modelers.

An increased number of silvicultural alternatives will be included. Intensive cultural practices such as planting genetically improved stock, controlling competing vegetation, applying fertilizers, and thinning need to be incorporated into growth and yield models for various species. Another silvicultural alternative--growing mixed species stands--also needs increased attention.

Estimates of future yield in traditional terms such as total cubic volume, while important, are inadequate. A complete analysis of management alternatives requires information on tree quality and product yields. Tree and wood properties must be related to silvicultural practices and this information incorporated into growth and yield models in order to make the models more realistic and useful.

Questions about forest response to environmental stresses will be asked with increasing frequency in the future. Growth and yield models that can provide insight into these questions will need to be more process oriented. The types of data to be gathered and model structure to be employed will need careful consideration.

CONCLUDING REMARKS

This brief summary of the status and future of growth and yield modeling has revolved around three key elements of data, analysis, and computation. It is important to be mindful of the interdependence of these elements in the quest for progress in growth and yield modeling. Analytical techniques require certain data characteristics and computational capabilities. Advances in computational technology allow the handling of data and application of analytical methods that previously were not feasible. More sophisticated analyses and increased computational power provide insight into data needs. Progress in the various elements is never going to be entirely smooth and even; consequently, progress in growth and yield modeling will wax and wane. The past has been a period of continual refinement and improvement of growth and yield modeling efforts, but the future promises even more rapid advancement.

While this presentation has necessarily been brief, more detailed information is available. The books by Avery and Burkhart (1983), Clutter et al. (1983) and Davis and Johnson (1987) all contain summary information on growth and yield modeling methods that are currently being used. A seminar series and workshop held at the University of Washington over the period January-March 1987 delved into current issues and future prospects of forest growth and yield modeling. Proceedings of this indepth analysis of the status and future of growth and yield prediction are contained in a volume edited by Chappell and Maguire (1987).

REFERENCES

Avery, T. E. and Burkhart, H. E., 1983: Forest measurements, 3rd Ed. McGraw-Hill Book Co., New York, NY, 331 pp.

Chappell, H. N. and Maguire, D. A. (Ed.), 1987: Predicting forest growth and yield: current issues, future prospects. College of Forest Resources, Univ. of Washington, Institute of Forest Resources Contribution No. 58, 95 pp.

Clutter, J. L., Fortson, J. C., Pienaar, L. V., Brister, G. H. and Bailey, R. L., 1983: Timber management: a quantitative approach. John Wiley & Sons, New York, NY, 333 pp.

Daniels, R. F. and Burkhart, H. E., 1988: An integrated system of forest stand models. Forest Ecol. and Manage., vol. 23, no. 2/3, pp. 159-177.

Davis, L. S. and Johnson, K. N., 1987: Forest management, 3rd Ed. McGraw-Hill Book Co., New York, NY, 790 pp.

Ford, E. D., 1984: The dynamics of plantation growth. In: Nutrition of Plantation Forests (Ed. G. D. Bowen and E. K. S. Nambiar), Academic Press, London, pp. 17-52.

Pool, R., 1989: Chaos theory: how big an advance? Science, vol. 245, no. 4913, pp. 26-28.

Smith, N. J. and Hann, D. W., 1984: A new analytical model based on the -3/2 power rule of self-thinning. Can. J. For. Res., vol. 14, no. 5, pp. 605-609.

INTERFACING FOREST INVENTORY WITH GROWTH AND YIELD MODELS

Alan R. Ek, Professor
Department of Forest Resources
University of Minnesota
St. Paul, MN 55108

ABSTRACT

Forest inventory and growth model design are viewed from the standpoint of facilitating use of forest growth models. Common uses of growth models are reviewed and suggestions are made for designing inventories to facilitate the development and use of such models. Alternatively, growth model design choices are considered as a means of taking advantage of forest inventory data types. Emphasis is placed on the completeness of data collection, growth model form and the packaging of growth models.

INTRODUCTION

Common uses of forest growth models are 1) to update forest inventories, 2) for resource analysis, particularly to assess silvicultural opportunities and to develop forestwide harvest schedules, and 3) for stand level "what if" experimentation. The overall usage of models for these purposes is low, yet much interest is expressed in such activity by managers and scientists (Walters 1989). In this paper we review the use of growth and yield models from the standpoint of facilitating their use. Opinion is offered as to why this use is low and suggestions are given to enhance usage.

We can take a number of steps to facilitate the development and use of forest growth models while we are designing forest inventories. There is much discussion of this in management organizations, but relatively little action occurs in practice. This is attributable to the fact that analyses are not often contemplated in detail prior to an inventory. The situation is exacerbated by the fact that few inventories are conceived and planned by those with specialized expertise in inventory design. Many resource managers may face only one such design effort in their career. Also, only recently has the idea of an inventory update with growth models been viewed as an alternative to an inventory. However, in many cases inventory designers are simply not clear as to what the needs might be for developing and using growth models. Growth modelers or analysts in turn are often unaware of the sampling context in which the use of their models is actually to be applied.

The sampling context for growth model applications implies use of the inventory data base as a basis for projection whether it be for an update or to project into the future. Lacking that, the model projections are made from a poorly defined point in space and time and have little value.

There are real urgencies for the development and application of growth models. These are for inventory update, treatment opportunity analysis for timber resource analysis and to address long term research questions concerning acid deposition or climate change impacts. We have done very little to address these research questions. Consequently, the potential impacts have been developed by simplistic speculative efforts that tend to ignore the enormous impact of initial forest conditions and the inertia of such ecosystems and the important role of ongoing management. From a science standpoint we need to do much better.

INVENTORY DESIGN

Inventories need to collect data periodically to facilitate the development of more, preferably more than one, models. The type of inventory is in part dependent upon the model forms to be considered. In fact, the model form might drive the inventory approach or vice-versa. If we are dealing with models of growth or change, then we should have a high interest in permanent plot installation and measurement. There are sound statistical reasons for that. Plots, in turn, need to be well distributed, with respect to the stand treatment conditions and other parameters (age, site, stand history, etc.) in the forest (Ek and Burkhart, 1976). Recall that inventories are often based on two phase sampling designs (Cochran, 1977). The growth models are, in effect, functions developed from the usual second phase of an inventory and applied back to the first phase data to make inferences about the whole population. Where data upon which models are based are drawn from specialized or atypical samples, inference capability can be significantly compromised.

At the very least we need to consider the development of models of an integral form. This means description of yield over age relationships. Appropriate jargon here is the collection of "cross-sectional" data where we describe forest stands of different age classes. We then assume through curve fitting that this pattern of development represents what we might expect if in fact we had observed individual stand development throughout a rotation. Empirical yield tables are an example of these yield over age relationships.

Don't throw away your old inventory data! Even temporary plot data from older inventories can be useful if the stratification or stand mapping criteria have been consistent over time. For example, old plot records describing the mean per acre characteristics of jack pine pole-sized stands of medium density on site index 60 are nearly as useful as new inventory plots in the same classification (Hansen and Burk,1985). The means of such tightly defined strata are unlikely to change much over time and the old data can effectively bolster the sample size for the development of simple models that describe yield as a function of age and site quality. This also argues for consistency in stratification in inventory design.

It is also wise to consider changes in forest type and composition as essential parts of growth and yield data collection. Briefly, we need to know the rates at which cover types change to other types through growth, harvest, other man-induced, or natural factors. Such observations are the basis for predicting future cover type change accurately, especially for regeneration following harvest. This implies the need for permanent plot observations and/or maintenance of detailed stand history information.

On a practical level, inventories today need to describe the trees and site in some detail. This means all trees one-inch Dbh and larger, or 4.5 feet tall and larger. Plot size is less important than the small tree size class detail. These small trees are eventual ingrowth and are essential to long term projection accuracy. If the development of true stand based growth models or individual tree based growth models is contemplated, then fixed size permanent plots are highly recommended. It is clear that much in the way of growth and change information can be developed from repeated measurements of Bitterlich horizontal point sample plots, but such methodology carries with it unnecessary complexity. Frankly, most management organizations do not have enough staff these days to deal with avoidable complexity.

Contrary to popular belief, we may not necessarily need a great deal of detail on the physiological characteristics of trees and/or site description information to make projections with physiological process models. Research by Sievanen et al. (1988) on parameter estimation for such models suggests that even very complex models may be estimated and applied to rather ordinary forest measurement data. However, it will become increasingly important to develop a small set of carefully measured permanent plots, well-documented and preferably near weather stations. I would suggest that each state should have a minimum of 30 such installations. Developing networks of that type across a region could provide enormous scientific value at low cost to treat long-term ecological questions. If these plots were part of a larger inventory, it would further help focus research on a representative sample of forest lands that have a clear sampling and inference context.

GROWTH MODELLING CONSIDERATIONS

As noted earlier the possible form of the growth model can be heavily dependent on detail of the forest inventory data. The amazing fact is that we have largely ignored some of the most available and most useful data for forest growth modelling for several decades. Early forest survey and growth modelling efforts from 1920 to 1950 developed empirical yield tables describing stand yield and to some extent approximations of growth with respect to age. All forest inventories collect the basic information for such models. The compiled stand characteristics and yield by age and site class are, in effect, a form of yield table. We can express those relationships conveniently today in model form to describe these average growth patterns. We have made important progress with that in the Great Lakes region as part of the Great Lakes Forest Growth and Yield Cooperative research (see Walters and Ek, 1989). We would profit by seeing such methodology applied to virtually all inventories. These yield tables become the simplest and perhaps most realistic characterization of forest development for natural stands and many types of managed stands. The construction of these tables is straightforward because we can observe the pattern of the data easily. They can also become the basis in part for judging the quality and utility of more complex models. It is disconcerting that we have developed few such empirical yield tables over the last several decades.

As permanent plot data is collected we can also augment those empirical models by growth expressions to describe changes and stand development upon management. Such models, if they are multispecies, facilitate applications with inventory data. For example, with an individual tree model, the model itself can use raw inventory data in tree list form to project a new future inventory in tree list form - and we can analyze it with the same programs developed for the inventory itself.

A decade ago "modelling" was not widely accepted. Managers viewed computer models with suspicion. Today modelling is a common term. Perhaps now we treat it too casually. In developing growth and yield models it is essential that we develop flexible and friendly growth implementation packages (GIP's). Given the lean staffing of management organizations today, failure to do so will mean the models will simply not be used. We need to integrate the models into common microcomputer data base management systems.

Regeneration Models

One element we have consistently ignored in research has been the development of regeneration models. There is relatively little literature on that. However, some effective regeneration models are also very easy to develop. We can use the inventory data from very young stands to characterize the regeneration following harvesting of stands. Simply put, when a aspen stand is harvested it is invariably replaced by a new aspen stand. In our modelling we have used inventory averages describing 15-year-old aspen stands to substitute or estimate the new stand following harvesting. Alternatively,if management dictates conversion to another cover type, we will replace the aspen standwith a 15-year-old stand from the inventory data describing the other type. This is one of the least biased and simplest forms of regeneration models. It is especially useful in describing the full species and size class mix of future stands. That detail is particularly important to successful use of individual tree growth models. As noted earlier, permanent plots and/or detailed stand history records are crucial to the development of accurate regeneration models.

Validation of Models and Uses

We must not forget that the inventory is the key data base for making updates and projections. The inventory data sorted and graphed in many ways is also the best test of our growth model with respect to stand age, site quality and treatment. That validation is also a source of direction with respect to the needs to localize models for geography, stand conditions, and treatment history. This suggests that we need to spend more time developing models from large inventory data bases as opposed to our current heavy concentration on small and specialized research data bases. For example, in my region there are thousands of plot records from statewide forest inventory efforts (see Hahn and Hansen, 1985) that can be put to good use now via simple models. At the same time such an effort would solidify and help provide direction for research on more complicated models.

Confusion in Applying Growth Models

Two situations exist. The first is illustrated by the multitude of growth models for loblolly pine in the southern U.S. Each model has been developed by different projects under slightly different assumptions about management. Another example is the considerable number of models for red and jack pine in the Great Lakes region of the U.S. and Canada. It becomes so difficult to choose one model that the end result is often that none are chosen. While some variables or parameters may be common to all the models, invariably there are enough differences that users need to restructure their inventory either in terms of field measurements or as it is stored on the computer.

The second situation that exists is where there are none or only a few models to choose from. This is typified in the Great Lakes region by the lack of simple models predicting growth and yield from mixed species stands. Over a large region, resource analysis may require growth and yield information for many species, many products and often very diverse climates. While it is probably not feasible to make one general model that treats all products and species, a package that implements a number of models can go a long way toward meeting needs. Whether there are ten models for a single species or ten models for different species, forest management staff do not want to maintain ten models with different input requirements. Such a situation leads to none being used in practice.

Integrating growth and yield models into a data base management system provides several obvious benefits. The first is that selecting, sorting, modifying, etc. of data for model usage can be done with standard queries. Second, incorporating models into management information system data bases facilitates ties to developing geographic information systems. If we do not take this step now, ten years from now people will have nice maps, but they still will not be able to project growth on the forest. Third, the tedious and time consuming part of writing a growth simulator is drastically reduced. Data base management system developers have probably already developed the basic screens and queries. Fourth, growth and yield model development and use of the models with operational data becomes more straightforward if they are all stored in a common data base.

DISCUSSION

In summary, we have much of the data and we have many of the analysis tools. We are just not making effective use of them. So now what? Each inventory design should specify a list of essentials necessary for developing meaningful growth and yield models and projections. Constructing the models should be considered part of the inventory data analysis effort. Further, that set of essentials should include an allowance for several types of growth models and how they might fit with desired inference capability and the data base management system under consideration.

Also, growth modelers need to consider the sampling context for application of models. It is invariably multiphase sampling with extrapolations from a model fit or developed from a subset of data to inferences for a broader level of sampling. If there is not a clear connection to an inventory in terms of the sampling context and the ready availability of predictor variables, the model development effort needs to be refined. Lacking that, inference capability will be severely limited. Additionally, start by developing simple models, including regeneration models, for the entire inventory. Given those, it is then appropriate to move to complex models

LITERATURE CITED

Cochran, W. G. 1977: Sampling techniques. 3rd edition. John Wiley and Sons. NewYork. 428 p.

Ek, A. R. 1987: Directions in forest growth and yield prediction. In: Predicting forest growth and yield: Current issues, Future prospects. H. N. Chappell and D. A.Maguire (eds.). University of Washington, College of Forest Resources, Instituteof Forest Resources, Contribution No. 58, pp. 31-43.

Ek, A. R. and H. E. Burkhart. 1976: Silvicultural considerations in timber inventory. Journal of Forestry 74:435-437.

Hahn, J. T. and M. H. Hansen. 1985: Data bases for forest inventory in the North Central Region. USDA Forest Service General Technical Report NC-101. 57 p.

Hansen, M. H. and T. E. Burk. 1985: Testing the consistency of forest cover type estimates across survey periods. In: Use of auxiliary information in natural resource inventories. VPI & SU, Blacksburg, VA. Society of American Foresters Pub. No. 86-01. pp 62-73.

Sievanen, R., T. E. Burk and A. R. Ek. 1988: Parameter estimation in a photosynthesis-based growth model. In: Forest growth modelling and prediction, Proceedings of IUFRO Conference. A. R. Ek, S. R. Shifley and T. E. Burk (eds.). USDA Forest Service General Technical Report NC-120, 2 vols., 1153 p. pp. 345-352.

Walters, D. K. 1989: An assessment of growth and yield research and education priorities as determined by a survey of GLFGYC members. University of Minnesota, Department of Forest Resources, Great Lakes Forest Growth and Yield Cooperative Pub. No. 4. 15 p.

Walters, D. K. and A. R. Ek. 1989: Development and application of a state-wide empirical growth and yield model for natural aspen stands. In: Proceedings of Aspen Symposium, July 25-27, 1989, Duluth, MN. In press.

Modelling with Data of the First Swiss National Forest Inventory

Edgar Kaufmann
Swiss Federal Research Institute for
Forest, Snow and Landscape
CH-8903 Birmensdorf, Switzerland

ABSTRACT

In connection with the preparatory work for the second Swiss National Forest Inventory, a model is being constructed with the aim of screening sampling designs and procedures to be applied. Some basic distributions, such as the distribution of tree diameter within a stand, are being derived from plot data of the first inventory. This paper describes the construction of a simple system of diameter distribution in uniform high forest. The observed distributions are fitted by means of the Weibull distribution, which employs three parameters. Because of the many-storeyed structure of many of the stands, the distributions were irregular in form. Dividing them into one distribution for trees of the upper storey and one for those of lower storeys allowed fitting by means of the Weibull distribution. The data were grouped into classes of stem count and basal area. All plots belonging to a particular class were grouped together for the fitting of a Weibull curve using the method of maximum likelihood. With a multiple linear regression over all classes within one stage of development, using the Weibull parameters as dependent variables, it was possible to construct a simple model with the Weibull parameters as linear functions of stand density and basal area. Other stand or site features exerted no substantial effect on the characteristics of the distributions. The aim of the diameter distribution system is to elaborate diameter distributions for each forest type and stage of development and for any given combination of stand density (number of trees per ha) and basal area (m2/ha) in such a way that they reflect the character of the forest type concerned.

INTRODUCTION

The first Swiss National Forest Inventory was conducted between 1983 and 1985. It was conceived as a multipurpose inventory and furnished basic data for the determination of the condition of the entire forest area of some 1.2 million ha. One of its major aims was to determine the standing volume. The terrestrial samples comprised some 11 000 permanent, circular plots distributed over a 1 X 1 km grid covering the whole country. The major areas for which conclusions were drawn were the five regions into which the country is divided for purposes of other fields of forest statistics; this division is based mainly on orographic features. These regions differ considerably in terms of production conditions, and in the first NFI varied greatly in the time-consumption, and consequently the costs, for terrestrial surveys, as the sample plots in mountainous regions were often difficult of access.

AIM OF THE MODEL

The first subsequent inventory, with emphasis on determining increment, is planned for the early 90's. As part of the preparatory work, a model is being developed to serve as one of the many aids in preliminary decision-making on changes in the inventory design and the evaluation of any new procedures which may come into question. The aim is to reduce the time-consumption and costs which pilot inventories involve. The plan is to use computer simulations of stands with the major objective of testing methods for estimating standing volume and increment.

The distributions of some important parameters are approximated to situations likely to occur in the major forest types. This involves consideration of regional differences. Data from the first NFI are being employed in the construction of a diameter distribution system and in studies on the spatial distribution of the trees. For the simulation of growth and yield, models developed in a National Research Programme for other purposes have been taken over (Lemm). The programme simulating increment is based on single-tree data. Consequently the diameter distributions derived from the sample plots of the NFI are not being projected.

THE DIAMETER DISTRIBUTION SYSTEM

This next section describes the first step in the construction of the model, i.e., the construction of the diameter distribution system. The discussion is limited to uniform high forest; analogous procedures are being employed for other forest types.

Material

Uniform high forest is characterized by the fact that most of the trees in each stand can be allocated to the same stage of development (rough age classification on the basis of DBHdom). This broad definition permits the classification of very heterogeneous forest structures under the heading of one forest type. Most of the stands have developed from regeneration plots and comprise trees of similar age.

Different forest types assume different degrees of importance in the various parts of the country; for instance, non-uniform forest in the Alps themselves, or sprout seedling and low or coppice forest on the southern slopes of Alps. In all regions, however, uniform high forest occupies the greatest area. In order to fit curves, all plots with forest or stand margins were excluded. Due to the small size of the stands here, this meant excluding some 30% of the plots. That left some 3850 plots with a total of 54650 diameter measurements available for the fitting of curves for the polewood and older stages in uniform high forest. The sample plots consist of two concentric circles covering 2 and 5 ares respectively. In the smaller plots, all trees with a DBH of at least 12 cm, were registered, in the larger all those with a DBH of 36 cm or more. That furnished a greater selection probability for the thicker stems. As no stem cores were taken, it was usually not possible to determine the exact age of the stands. Consequently, the stage of development, as estimated by the field crews through purely visual observation, was taken as a rough measure for age classification. A diameter distribution system was elaborated for each stage of development.

Estimation of the Weibull Parameters

For the description of the diameter distributions the Weibull distribution, which is based on three parameters, was chosen. In recent years, this particular distribution has been widely used for these purposes because it is flexible and can be easily integrated. It allows the construction of a whole range of reversed J-shaped and mound-shaped curves. The probability density function is given by (Harter and Moore 1965, Bailey and Dell 1973):

$$f(x) = (c/b) \ \ ((x-a)/b)^{c-1} \ \ \exp(-(x-a)/b)^{c})$$

and the cumulative densitiy function by:

$$F(x) = 1 - \exp(-((x-a)/b)^{c})$$

where a = location parameter (can be interpreted as the smallest possible diameter);
b = scale parameter
c = shape parameter

The three parameters, a, b, and c, were estimated through the method of maximum likelihood (ML). They were obtained by equating the partial derivatives of the logarithm of the likelihood function with respect to the parameters to zero. The equations for the terms a and c had to be solved by means of iterative procedures. Their ML equations are (Harter and Moore 1965):

$$\text{a:}\quad (1-c)\sum_{i=1}^{n}(x_i-a)^{-1} + cb^{-c}\sum_{i=1}^{n}(x_i-a)^{c-1} = 0$$

$$\text{c:}\quad n(1/c-\ln(b)) + \sum_{i=1}^{n}\ln(x_i-a) - \sum_{i=1}^{n}((x_i-a)/b)^{c}\ln((x_i-a)/b) = 0$$

The term b can be estimated with every iteration through the following equation (Bailey and Dell 1973):

$$b = ((1/n)\sum_{i=1}^{n}(x_i-a)^{c})^{1/c}$$

The variance-covariance matrix of the likelihood estimators was computed with the classic methods. That is, by inverting an information matrix whose elements are the negatives of the expected values of the second order derivatives of logarithms of the likelihood function with respect to the parameters (Cohen 1965, Linder and Berchtold 1982).

In all computations, the different selection probabilities of the individual trees were taken into account through appropriate weighting. Verification with data stemming solely from the smaller of the two circular plots revealed no bias in comparison with the results from calculations based on data from all trees in both plots.

Procedure and Results

The distributions of diameter are very irregular, especially in the older stages of development (Fig. 1, curve I), and cannot, in this form, be fitted with the Weibull distribution without recourse to fairly complicated types of the distribution (see e.g. Cao and Burkhart 1984, Zutter et al. 1986). Division into two distributions, one for the trees of the upper storey and one for those of the middle and lower storeys, produced a mound-shaped curve for the upper storey and a reversed J-shaped curve for the lower ones. These two distributions were suitable for fitting with the Weibull distribution. As the trees of the lower storeys may comprise a considerable stem count they may not be neglected in models of diameter distributions. The irregular distributions primarily reflect the fact that for silvicultural reasons the stands were purposely cultivated in a multi-storeyed form.

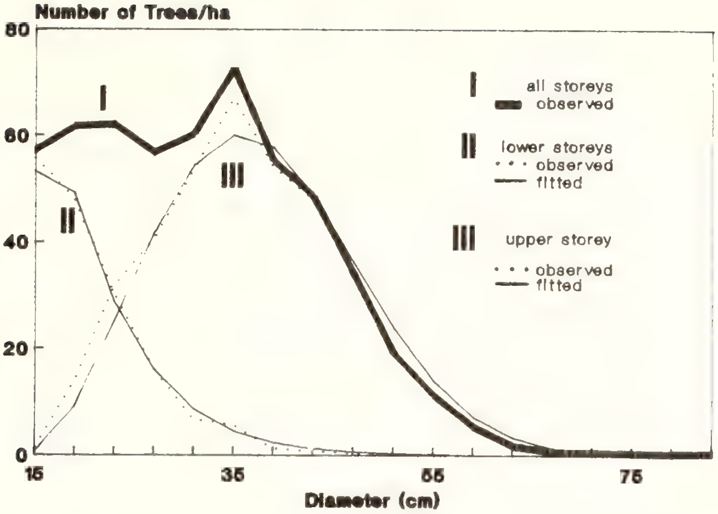


Fig. 1. Diameter distributions development stage mature timber II: Stem count class: 450-650 trees/ha, basal area class 40-50 m2/ha. Observed data NFI and fitted Weibull distributions of upper storey and lower storeys.

The input parameters needed for the models comprise one forest type, one stage of development, and the number and size of the stands to be generated. For any given strata, stand density (i.e., number of trees per ha) and basal area (m2/ha) are randomly selected with correlations and density distributions which have been approximated to values obtained in the inventory data. After generation of a spatial point pattern, the individual trees are allocated a diameter. The aim of the diameter distribution system is to spread any given reasonable combination of stem count per ha (SCPH) and basal area (BASA) over the individual trees in such a way that a generated stand exhibits the pre-set parameters and the distribution of diameter approximates that actually found in reality. In our forests, stand density and basal area of a particular development stage can vary greatly. One major reason for this lies in the silvicultural techniques employed: there is little conformity, and though thinning is conducted on the same principles, the intensity of operations depends greatly on the personal judgement of the forester in charge.

Consequently, the parameters of the Weibull distribution are defined as functions of stand density and basal area for each stage of development. In order to have an adequate number of trees for the estimation of the Weibull parameters in each case, stem count and basal area are divided into classes. All sample plots are allocated to the class they represent, irrespective of their geographical location. In each case the data set for the computation of a Weibull curve comprised all trees of all sample plots in the given class. Table 1 illustrates the procedure, using the stem count class 450-650 trees/ha for the upper storey in the development stage mature timber II (estimated DBHdom 41 -50 cm). As the three parameters are closely correlated with each other, so that any change in one affects the other two, the procedure was conducted in the following steps (Table 1):

Table 1. Estimation of Weibull Parameters: Development stage mature timber II, upper storey, stem count class: 450-650 trees/ha, basel area classes 30-40 m2/ha up to 50-60 m2/ha (standard deviations of maximum likelihood estimates in brackets, values determined by regression analysis underlined, n = number of trees).

Basal Area m2/ha	Step	Weibull - Parameter a	b	c
30-40	1	12.27 (0.58)	23.70 (0.80)	2.27 (0.10)
	2	12.02	23.99 (0.20)	2.30 (0.005)
(n=601)	3		24.11 (0.17)	2.36
	4		24.40	
40-50	1	11.20 (0.45)	27.05 (0.60)	2.60 (0.08)
	2	13.10	25.01 (0.11)	2.40 (0.003)
(n=1090)	3		25.10 (0.08)	2.45
	4		25.24	
50-60	1	14.04 (0.61)	25.35 (0.77)	2.51 (0.10)
	2	14.18	25.20 (0.13)	2.50 (0.004)
(n=863)	3		25.28 (0.12)	2.54
	4		26.08	

1. simultaneous estimation of all three parameters (method ML) and their variance in each class;
2. determination of parameter a through linear regression over all classes using weighted least squares;
   dependent: parameter a
   independent: SCPH, BASA (class means)
   weights: inverses of the variances of the estimates of parameter a
   estimation of c (method ML) with the given a;
3. determination of c through linear regression
   estimation of b with ML using the given a and c;
4. determination of b through linear regression.

This procedure produces a simple linear model, for instance for the case mature timber II, upper storey:

$$
\begin{aligned}
a &= 10.88 + 0.108\ \mathrm{BASA} - 0.0048\ \mathrm{SCPH} \\
b &= 31.15 + 0.084\ \mathrm{BASA} - 0.0176\ \mathrm{SCPH} \\
c &= 2.89 + 0.009\ \mathrm{BASA} - 0.0015\ \mathrm{SCPH}. \qquad (1)
\end{aligned}
$$

where BASA = basal area m2/ha, SCPH = stem count/ha

Given constant stem count, the mean diameter increases with increasing basal area (Fig. 2), and given constant basal area decreases with increasing stem count (Fig. 3). This relationship can be adequately described through a linear change in the three parameters of the Weibull distribution.

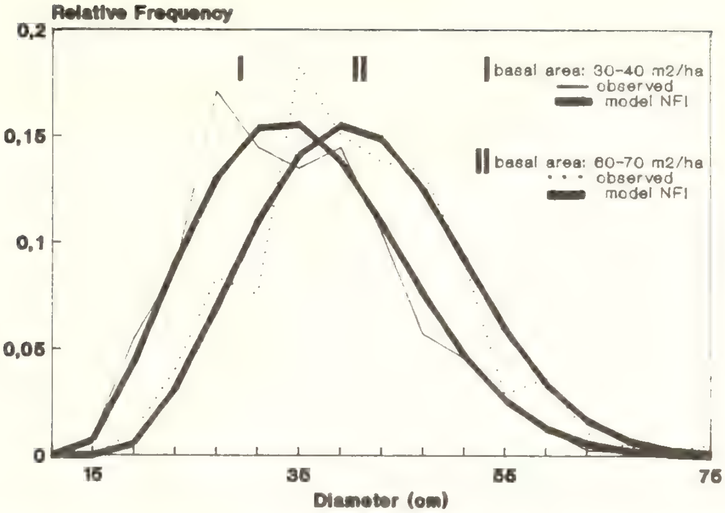


Fig. 2. Diameter distributions development stage mature timber II, upper storey, stem count class 450-650 trees/ha. Observed data NFI and distribution model NFI (1) for basal area class 30-40 m2/ha and class 60-70 m2/ha.

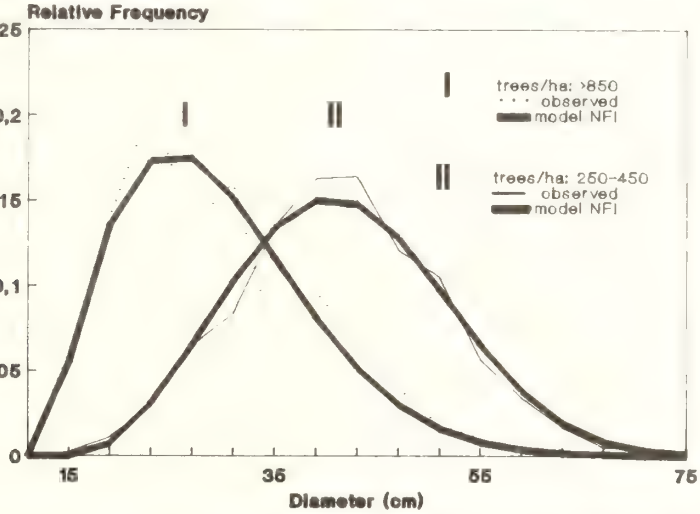


Fig. 3. Diamater distributions development stage mature timber II, upper storey, basal area class 40-50 m2/ha. Observed data NFI and distribution model NFI (1) for stem count class 250-450 trees/ha and class >850 trees/ha.

The regions vary greatly as regards site conditions, intensity of management, and species composition of the stands. In the flat Mittelland two thirds of the forest, in terms of area, stands on very good sites, but only 3% in the mountains. Here, the forest predominantly consists of purely coniferous stands, while in the more equable Jura purely deciduous and mixed forest dominate. In almost 80% of the forests in the Mittelland, wood has been harvested within the past ten years, whereas in more than 70% of the forests on the southern slopes of Alps and some 45% of the mountain forests no harvesting has been conducted in the past 20 years.

In order to take any substantial influence of these differences on the diameter distributions into consideration, the Weibull parameters for every single plot containing more than 10 trees were estimated and subjected to stepwise regression analysis. For each run the dependent variable was one of the Weibull parameters, while the independent variables were mean basal area diameter, proportion of conifers, site quality, number of years since the last harvest, slope inclination, and elevation. No important correlations were found for parameter a. For parameters b and c, the mean basal area diameter was entered in each run as the first variable in the regression and comprised by far the greatest proportion in the multiple r2. The only other factor to contribute any considerable effect on the r2 of the shape parameter was the proportion of conifers.

Several diameter distribution systems were elaborated with the procedure described above. Subsequently, in addition to division into classes of stem count and basal area, the data were classified according to other criteria: high or low proportion of conifers; very good or very poor site conditions; intensive or lax management; location of the sample plot in the mountains or the Mittelland. This produced two independent models for comparison in each case. Visual inspection of the curves showed that the differences are obviously too small to require consideration in the evolution of diameter distributions. Samples of varying size were generated from the fitted Weibull distributions for the class with the highest frequency (SCPH 450-650, BASA 40-50 m2) for trees of the upper storey belonging to the development stage mature timber II. Pairs of samples from two independent models were subjected to the Kolmogorov-Smirnov two sample test. In the majority of cases they were only rejected at the 5% level when they comprised several hundred elements.

With each successive division of the data the distributions become more and more irregular and the estimates of the parameters correspondingly less accurate. Consequently, differences may arise from the way in which the model is constructed and not from genuine differences in the actual data.

Comparison with yield study sample plots

Combined classes for stand density and basal area were compiled from a large number of small sample plots which together formed a heterogeneous aggregate with respect to many parameters. It seemed probable that the variances of these distributions could not be representative for individual stands. The model was therefore compared with 4 arbitrarily chosen yield sample plots. These plots are independent of and larger than the NFI plots (0.3 -0.5 ha). They had been regularly and intensively managed since their establishment. The diameter distributions employed were those pertaining immediately before any silvicultural operation. The variation coefficients of the fitted curves of two of the plots, in which an extended, schematized type of thinning had formerly been applied, were all about 40% smaller than those of the corresponding NFI model. The two other plots had been subjected to the manner of thinning now generally used in Switzerland. The variation coefficients of the fitted Weibull distributions for these plots generally lay within the order of magnitude of the NFI model. Though these comparisons do not allow of generalizations, they do confirm that similar distributions evolved from aggregates of small sample plots occur in individual stands (see Fig. 4), though here only very small areas with homogeneous conditions are involved.

CONCLUSION

The diameter distribution system described above is one of the bases for the model now being elaborated. The aim of this model is to furnish an approximate picture of the different forest types of varying importance in the different regions of Switzerland through a few important distributions.

Only the great differences actually obtaining between the forest types are being considered. It is more important to elaborate an informative model with the simplest possible construction than to compute the most nearly accurate fitting of distributions. The model will be used to screen sampling designs and procedures for the second NFI survey.

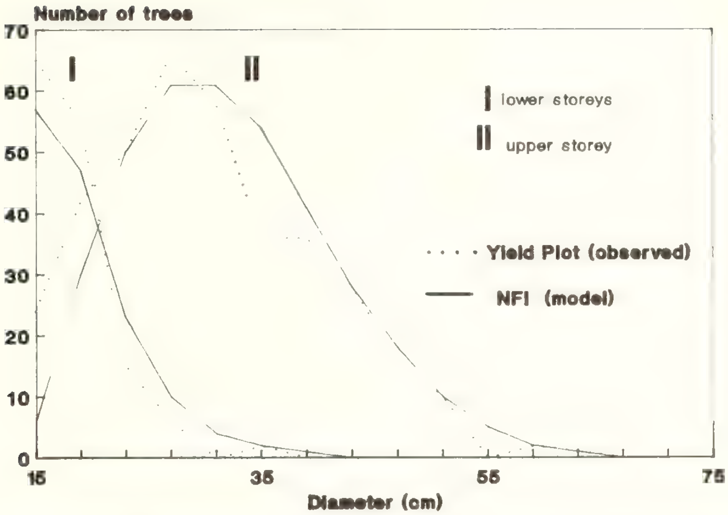


Fig. 4. Diameter distributions of upper storey and lower storeys of a yield study sample plot (861 trees/ha, basal area 52.04 m2/ha) compared with distribution model NFI (1).

LITERATURE

Bailey, R.L. & Dell, T.R., 1973: Quantifying diameter distributions with the Weibull function. Forest Sci. 19(2): 97-104.

Cao, Q.V. & Burkhart, H.E., 1984: A segmented distribution approach for modeling diameter frequency data. Forest Sci. 30(1): 129-137.

Cohen, A.C., 1965: Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples. Technometrics 13: 171-182.

Harter, H.L.& Moore, A.H., 1965: Maximum likelihood estimation of the parameters of gamma and Weibull populations from complete and from censored samples. Technometrics 7: 639-643.

Kilkki, P. & Päivinen, R., 1986: Weibull function in the estimation of the basal area dbh-distribution. Silva Fennica 20(2): 149-156.

Lemm, R.: Simulation von Wertverlusten durch Waldschäden. Proj. NFP12, not published.

Linder, A. & Berchtold, W., 1982: Statistische Methoden II. Birkhäuser, Basel, 295p.

Schweizerisches Landesforstinventar, 1988: Ergebnisse der Erstaufnahme 1982-1986. EAFV-Bericht Nr. 305.

Shiver, B.D., 1988: Sample sizes and estimation methods for the Weibull distribution for unthinned slash pine plantation diameter distributions. Forest Sci. 34(3): 809-814.

Zingg, A., 1988: Schweizerisches Landesforstinventar, Anleitung für die Erstaufnahme 1982-1986. EAFV-Bericht Nr. 304.

Zutter, B.R. et al., 1986: Characterizing diameter distributions with modified data types and forms of the Weibull distribution. Forest Sci. 32(1): 37-48.

Updating Methods for Forest Inventories

Moderator: Richard A. Birdsey

Richard Birdsey, Updating Methods for Forest Inventories - An Overview

Kurt B. Teuber, Remote Sensing Techniques for Updating Forest Inventories

W. Brad Smith, Inventory Updates: Overview of the North Central Experience

Colin D. MacLean and Neil McKay, Updating Inventories in Western Oregon and Washington Using Tree Growth, Mortality and Harvest Models Developed from Permanent Plot Data

Raymond L. Czaplewski, Kalman Filter to Update Forest Cover Estimates

Updating Methods for Forest Inventories - An Overview

Richard Birdsey
Forest Service - USDA
Washington, DC 20090-6090

ABSTRACT

Updating of forest inventory estimates is done because full-scale inventories are too expensive to be conducted as frequently as information is needed. There are many tools and techniques available for updating inventory estimates, including trend extrapolation and time-series analysis, transition probability matrices, timber supply models, growth and yield models, sampling methods, remote sensing and geographic information systems, and combinations of techniques. Because updating and monitoring are closely related, techniques already proven for updating will have applications for forest health monitoring.

THE NEED FOR UPDATING

The foundations for monitoring the forest resources are inventories that are designed to provide information about changes or trends. These successive inventories are said to be integrated temporally (Lund 1986). Inventory design and frequency determine how well the forest resource changes are monitored. In the U.S., a system of permanent sample plot locations is remeasured in a continuous cycle of approximately 10 years. The length of the cycle is related more to funding and logistics than to design considerations. Update frequency should ideally be related to the dynamics of change and the importance of changes (Milazzo 1981). Most data users feel that a 10-year cycle is too long, considering the rapid changes taking place in the forest and the current needs to monitor timber supplies, growth trends, and interactions between forests and the atmosphere.

The dynamics of Midsouth forests were illustrated in a recent survey of Louisiana (Rosson et. al. 1987). Land-use shifts between forest and nonforest uses involved 1.8 million acres in 10 years, roughly 13 percent of Louisiana's 13.9 million acres of timberland. During the same period, 1.6 million acres were clearcut, 2.4 million acres were partially cut, and 1.7 million acres were disturbed in some other way. Only 7.7 million acres or about 55% of the timberland were undisturbed between surveys.

When do estimates of forest parameters become outdated? This is a difficult question to answer for multiple purpose forest inventories because of the number of variables estimated and the many classifications used to post-stratify the data. Birdsey (1989) used analysis of variance techniques to test the hypothesis that there was no difference between estimates of mean volume, growth, and removals per acre for 1978 and 1988 inventories of a survey unit in Southwest Arkansas. Estimates were classified by 2 species groups, 2 ownership groups and 5 forest type groups. Results showed no difference for mean volume estimates, significant differences for mean growth estimates, and a significant difference for mean softwood removals but not hardwood removals. Area estimates by forest type were substantially different for the two surveys. These tests of a few of the many variables estimated showed that after 10 years a new inventory was indeed necessary at the survey unit aggregation, and would certainly be required at the county level.

Extensive forest inventories are expensive and time-consuming. Despite research in survey sampling and methodology, advanced computing and data-base management, and timely remote sensing, locating and measuring sample plots involves a lot of time and money. It seems unlikely that simple repetition of survey samples will ever be frequent enough to effectively monitor rapidly changing forests.

Another way to keep information current is to use some sort of updating procedure between remeasurement samples. Updating can be defined as as estimate of the current forest condition derived by modeling the dynamic change in a forest from a known time and condition in the past (Smith and Raile 1979). The key word in this definition is "modeling", which distinguishes an inventory update from a straightforward repetition of the previous inventory. Another key concept is that updating methods can be arrayed in a continuum of methods from simple to complex. The simplest updating method might be an accounting for land area changes or an extrapolation of past trends, while the most complex methods will involve combinations of techniques and relatively elaborate procedures. This overview begins with a description of simple, basic tools for updating and progresses to complex, integrated updating systems.

SOME BASIC TOOLS FOR UPDATING

Time series analysis is the classical method of forecasting based on the analysis of historical data and the extrapolation of past trends into the future. The basic assumptions are that trends of the past will continue through the forecast period without change, and that no unforeseen events will arise to disrupt the forecast. Trend extrapolation can involve a simple straight-line projection or may follow a more complex curvilinear form. Most fundamental statistical textbooks for business and economics describe forecasting and time series analysis methods in detail (for example, see Neter, Wasserman, and Whitmore 1973; Pfaffenberger and Patterson 1977).

Extrapolation of past trends can be a simple and effective technique for short-term updating of a forest inventory. Longer term extrapolation increases the probability that predictions based on past trends will deviate substantially from actual trends. Nevertheless, many users of periodic forest inventory data use extrapolation of past trends in volume, growth, and removals to estimate current volume from an inventory that is 5 or 10 years old. Even complex, integrated models use this technique for updating variables for which no reliable sub-model has been developed or applied (U.S. Department of Agriculture, Forest Service 1987).

A simple way to update volume by extrapolating past trends is to use the following equation:

$$V_2 = V_1 + n(G - R)$$

where V_1 = most recent volume estimate
V_2 = updated volume estimate
G = recent estimate of net annual growth
R = recent estimate of annual removals
n = number of years to update

This system could be made more elaborate by adjusting removals for recent harvest estimates from other data sources.

A basic tool for updating area estimates by any classification or updating a diameter-class distribution is the Markov transition probability matrix. Transition probability matrices show the proportion of a given class that shifts to another class between two points in time. Rows are defined as past class and columns as current class. If $\mathbf{T}$ is a transition probability matrix derived from 2 inventories spanning $\mathbf{k}$ years, and $\mathbf{A}_t$ a row vector representing the initial distribution of classes, then the class distribution at time $\mathbf{t+k}$ is $\mathbf{A}_t\mathbf{T}$. The class distribution after $\mathbf{n}$ intervals is $\mathbf{A}_t\mathbf{T}^n$. As defined here the system is limited to increments of $\mathbf{k}$ years, the intersurvey period, and is restricted by the assumption that transition probabilities are static. Both of these restrictions can be overcome, and for updating for shorter periods it is necessary to derive an annual probability matrix from a periodic one (Edwards et.al. 1985). Classification systems for area updating can include several variables such as forest type, ownership, and disturbance, but can require large data sets to develop precise estimates of transition probabilities for each cell in the transition matrix (Birdsey 1989).

The probability of an event occurring is used for modeling tree harvest or mortality in some growth and yield models, and is often estimated with regression techniques (Hamilton 1974). Event probabilities can also be used to model sample plot disturbance for updating.

MORE BASIC TOOLS: REMOTE SENSING, SAMPLING, AND GIS

Remote sensing is used in most large-scale inventories as the principal source of information on forest area, and in inventory updates as a source of information on area change. Remote sensing techniques may be used alone to update the condition of specific tract of land, or as part of some sample survey update.

Sampling designs for updating are usually dependent on the previous survey because of gains in efficiency over estimating change from two independent samples (Cunia 1985). With statistically independent samples, the difference between 2 random variables has a variance equal to the sum of the 2 variances, and the value of the variance is relatively high compared to the magnitude of the difference. If sample plots are remeasured, the successive values are usually highly correlated and the variance of the difference between 2 random variables equals the sum of their variance minus twice their covariance. Thus most contemporary continuous forest inventory designs involve remeasurement of some or all of a system of permanent sample plots. Basic sampling techniques for updating are similar to sampling techniques used for successive inventories.

GIS is rapidly emerging as the principal tool to relate various kinds of spatial information such as land cover, roads, or political boundaries. GIS is used extensively to overlay interpreted aerial photographs or satellite digital data with any mapped information, and has enormous potential for updating cover condition for specific tracts of land or sample points.

TIMBER SUPPLY MODELS FOR UPDATING

Econometric models use formal statistical procedures to relate variables that are difficult to forecast to variables that are thought to be more reliably predicted with other models. For example, forest acreage trends in the Southeast have been projected by modeling relationships among changes in population, income, socioeconomic variables, and forest acreage (Alig 1986).

The Timber Resource Analysis System (TRAS) has been used in past Forest Service assessments of timber supplies (Larson and Goforth 1974). TRAS is a stand-table projection model in which the timber resource is aggregated by ownership, by softwood and hardwood, and by 2-inch diameter classes. Projections, updates, or backdates are made by simulating annual changes in timberland area and numbers of trees.

The current inventory model used by the Forest Service for projecting timber supplies is called the Aggregate Timberland Assessment System (ATLAS). ATLAS is an area-based, yield-table system that projects acres by detailed strata for periods consistent with inventory stand-age classes. The inventory is represented by an array of acreage cells classified by region, ownership, management type, site productivity class, stocking class, management intensity, and age class. Acreage is shifted among the cells, and volume and growth estimates are derived from acreage movements through assumed yield tables. ATLAS can be used for updating, but in practice, simple interpolation and extrapolation are used to update the data to a common base year. ATLAS projections begin with inventory data from the most recent State survey. The first projection period is used to calibrate the ATLAS input data with the updated historical data.

UPDATING WITH GROWTH AND YIELD MODELS

The methods of growth and yield modeling are of interest because these models are often used for updating as well as projections, and the methodology is diverse and well-documented. This section briefly describes models of forest growth and yield developed by foresters for production planning, silvicultural research, and forest management.

Several authors have documented the wide array of growth and yield models available (Trimble and Shriner 1981; Alig et.al. 1984). The earliest models were no more than yield tables constructed from temporary sample plots (U.S. Department of Agriculture, Forest Service 1929). Subsequent models relate stand growth to variables such as age, height, and stand density; relate stand growth to stand structure (diameter distribution); or simulate growth, removals, and mortality at the individual tree level for aggregation to the stand level.

Growth and yield models are usually used to forecast the development of particular stands under various levels of treatment, but can be used in various ways to update stand inventories or project regional inventories (Alig et.al. 1984). The type of model selected should be commensurate with the specific estimate to be updated and the characteristics of the forest under consideration. Individual-tree simulation models are the most flexible in terms of their ability to model stands with different species composition and structure. A good example of this kind of model is the STEMS system developed for the North Central States and eventually applied for inventory updating (Belcher et.al. 1982; Smith and Hahn 1986).

APPROACHES TO UPDATING WITH MULTIPLE TECHNIQUES

Midcycle or interim surveys combine subsamples of existing plots and a regression estimator with estimates of area change from current remote sensing to update inventory statistics (Scott 1983, Thomas 1983). Using similar techniques, Maclean (1984) first stratified permanent sample plots into disturbed and undisturbed classes by photo interpretation, then subsampled the undisturbed plots for a regression update of all undisturbed plots. These procedures are more efficient than full remeasurement. Other sampling strategies may be even more efficient (Schreuder and Thomas 1985).

These examples of sampling for updating are special cases of the general procedure of Sampling with Partial Replacement (SPR) developed in the U.S. by Ware and Cunia (1962), with no temporary plots at the second measurement. SPR has gained worldwide acceptance as an efficient sampling design that minimizes the measurement of expensive permanent sample plots while producing an efficient estimate of change (Matis et.al. 1984).

Although sample-based updates can be very efficient, they still require a substantial investment of resources for design, sample selection, fieldwork, acquisition and interpretation of aerial photography, and data compilation and analysis. For frequent updating of many different state inventories, some other modeling approach may be more appropriate.

Combined sampling and modeling approaches can increase the efficiency of sampling when plots can be separated into disturbed and undisturbed classes by photointerpretation. Hahn (1984) described a procedure to model changes on undisturbed plots and sample disturbed plots. A few undisturbed plots are remeasured to produce a regression estimate for the whole undisturbed class. A similar approach by Fairweather and Turner (1983) used predictions derived from growth models as auxiliary variables in a double sampling framework. At least one timber products company uses very similar procedures to maintain a current inventory of their timberlands (Hasse and Robar 1983). In the North Central U.S., the STEMS model has been fully integrated into the regular periodic inventory sample as a means to update estimates for undisturbed sample plots based on a subsample (Spencer et.al. 1988).

Czaplewski et. al. (1987) have proposed combining estimates from growth and yield models, remote sensing, and regular forest inventories with the Kalman filter. The method uses a linear filter (Kalman filter) to achieve sampling efficiency (Dixon and Howitt 1979; Jameson 1983).

In a system designed for annual updating, Birdsey (1989) combined a matrix model to update areas by forest type with econometric and other techniques to incorporate exogenous data on harvesting and planting. Subsampling and change detection by remote sensing would add considerable validity to this combination of techniques.

CONCLUSION: UPDATING AND FOREST HEALTH MONITORING

Updating of national inventory statistics has been an issue for some time, although more frequent inventories in many parts of the country have recently lessened the need to update or conduct midcycle inventories. Still, for many purposes, a 10-year average national inventory cycle is too long. A recent issue requiring more frequent estimates is forest health monitoring, highlighted by public concern about the effects of acid rain, prospective climate changes, and intense forest pest outbreaks.

The Forest Inventory and Analysis (FIA) program of the U.S. Forest Service has monitored forest conditions for many years. The periodic inventories are the only source of comprehensive, long-term data about the Nation's forest lands. Most proposals for national forest health monitoring envision a role for the FIA program to collect additional data at more frequent intervals for monitoring purposes.

Because updating and monitoring are closely related, techniques already proven for updating will have applications for forest health monitoring as well. It is likely that any system eventually developed to detect forest health problems will involve many of the tools and techniques highlighted in this overview.

REFERENCES

Alig, Ralph J., 1986: Econometric analysis of the factors influencing forest acreage trends in the Southeast. For. Sci. 32(1): 119-134.

Alig, Ralph J.; Parks, Peter J.; Farrar, Robert M. Jr.; and
Vasievich, J. Michael, 1984: Regional timber yield and cost information for the South: modeling techniques. Gen. Tech. Rep. RM-112. Fort Collins, Colorado: U.S. Department of Agriculture, Forest Service, 28 p.

Birdsey, R.A., 1989: A dynamic matrix model for updating forest resource inventories. PhD Dissertation, State University of New York, College of Environmental Science and Forestry, 267 p.

Belcher, David W.; Holdaway, Margaret R.; and Brand, Gary J., 1982: A description of STEMS. Gen. Tech. Rep. NC-79. St. Paul, MN: U.S. Department of Agriculture, Forest Service, 18 p.

Cunia, Tiberius, 1985: More basic designs for survey sampling: two-stage and two-phase sampling, ratio and regression estimators and sampling on successive occasions. Forest Biometry Monograph No. 4, Prelim. Ed. Syracuse, NY: State University of New York, 368 p.

Czaplewski, R.L.; Alig, R.J.; and Cost, N.D., 1987: Monitoring land/forest cover using the Kalman filter: a proposal. In Proceedings, Forest Growth Modelling and Prediction, Vol. 2, Ed. by Ek. et.al., Minneapolis, MN: Society of American Foresters, 1089-1096.

Dixon, Bruce L.; and Howitt, Richard E, 1979: Continuous forest inventory using a linear filter. For. Sci. 25(4): 675-689.

Edwards, Clark; Smith, Matthew G.; and Peterson, R. Neal, 1985: The changing distribution of farms by size: A Markov analysis. Ag. Econ. Research; 37(4): 1-16.

Fairweather, Stephen E. and Turner, Brian J, 1983: The use of simulated measurements in double sampling for successive forest inventory. In: Proceedings, Renewable Resource Inventories for Monitoring Changes and Trends; 1983 August 15-19, Corvallis, Oregon. Corvallis, OR; Oregon State University; 609-612.

Hahn, Jerold T., 1984: Forest inventory in the North Central Region - data bases, sampling procedures and data processing. In: Lund, H. Gyde, ed. Preparing for the 21st century: Proceedings of the forest land inventory workshop; 1984 March 26-30; Denver, CO. Washington, DC: U.S. Department of Agriculture, Forest Service, Timber Management Staff; 86-93.

Hamilton, David A., 1974: Event probabilities estimated by regression. Res. Pap. INT-152. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 20 p.

Hasse, William; and Robar, Stephen, 1983: Keeping a current inventory on Champion International lake states timberland. In: Proceedings, Renewable Resource Inventories for Monitoring Changes and Trends; 1983 August 15-19, Corvallis, Oregon. Corvallis, OR Oregon State University, 120-123.

Jameson, Donald A., 1983: Sampling efficiency for monitoring of model based forest plans. In: Proceedings, Renewable Resource Inventories for Monitoring Changes and Trends; 1983 August 15-19, Corvallis, Oregon. Corvallis, OR; Oregon State University, 568-572.

Larson, R.W.; Goforth, M.H., 1974: TRAS: a timber volume projection model. Tech. Bull. 1508. U.S. Department of Agriculture, Forest Service, Washington, DC., 15 p.

Lund, H. Gyde, 1986: A primer on integrating resource inventories. General Technical Report WO-49. Washington, D.C.: U.S. Department of Agriculture, Forest Service, 64 p.

MacLean, Colin D., 1984: Regression updates of inventory plots -- a low cost alternative to remeasurement. In: Proceedings, Inventorying Forest and Other Vegetation of the High Latitude and High Altitude Regions; 1984 July 23-26; Fairbanks, Alaska, 68-73.

Matis, K.G.; Hetherington, J.C.; and Kassak, J.Y., 1984: Sampling with partial replacement - a literature review. Commonw. For. Rev. 63(3): 193-206.

Milazzo, Valerie A., 1981: Considerations in planning a land use and land cover resources inventory update program. In: Proceedings, In-Place Resource Inventories: Principles and Practices; 1981 August 9-14; Orono, Maine. Orono, ME; University of Maine, 135-143.

Neter, John; Wasserman, Williams and Whitmore, G.H., 1973: Fundamental statistics for business and economics. 4th ed. Boston, MA: Allyn and Bacon, Inc., 781 p.

Pfaffenberger, Roger C.; and Patterson, James H., 1977: Statistical methods for business and economics. Homewood, IL: Richard D. Irwin, Inc., 750 p.

Rosson, James F. Jr.; Rudis, Victor A.; McWilliams, William H.; and Frey, Paul, 1987: Forest resources of Louisiana. Resource Bulletin SO - 130. New Orleans, LA: U.S. Department of Agriculture, Forest Service, 81 p.

Schreuder H.T.; and Thomas, C.E., 1985: Efficient sampling techniques for timber sale surveys and inventory updates. For. Sci. 31(4): 857-866.

Scott, Charles T., 1983: Example of midcycle updating to assess catastrophic change. In: Proceedings, Renewable Resource Inventories for Monitoring Changes and Trends; 1983 August 15-19, Corvallis, Oregon. Corvallis, OR; Oregon State University, 555-557.

Spencer, John S. Jr.; Smith, W. Brad; Hahn, Jerold T.; and Raile, Gerhard K., 1988. Wisconsin's fourth forest inventory. Resource Bull. NC-107. St. Paul, MN: U.S. Department of Agriculture, Forest Service, 158 p.

Smith, W. Brad and Hahn, Jerold T., 1986: Michigan's forest statistics, 1987: An inventory update. Gen. Tech. Rep. NC-112: St. Paul, MN, U.S. Department of Agriculture, Forest Service, 44 p.

Smith, William B.; and Raile, Gerhard H., 1979: FREP: Application of the tree growth projection system for inventory. In: Proceedings, Forest Resource Inventories Vol. 1, 1979 July 23-26, Fort Collins, Colorado. Fort Collins, CO; Colorado State University, 223-230.

Thomas, Charles E., 1983: Maintaining mid-cycle inventories of forest resources. In: Proceedings, Renewable Resource Inventories for Monitoring Changes and Trends; 1983 August 15-19, Corvallis, Oregon. Corvallis, OR; Oregon State University, 386-390.

Trimble, Judy L.: and Shriner, Carole L., 1981: Inventory of United States forest growth models. Report prepared for: Oak Ridge, Tennessee; Oak Ridge National Laboratory, 114 p.

U.S. Department of Agriculture, Forest Service, 1929: Volume, yield, and stand tables for second growth southern pines. Misc. Pub. 50. Washington, D.C., 202 p.

U.S. Department of Agriculture, Forest Service, 1987: The South's fourth forest-alternatives for the future. For. Resource Rep. No. 24. Govt. Print. Off., Washington, D.C. 512 p.

Ware, K.D.; and Cunia, T., 1962: Continuous forest inventory with partial replacement of samples. Forest Science Monograph 3, 40 p.

Remote Sensing Techniques for Updating Forest Inventories

Kurt B. Teuber
Research Forester
USDA Forest Service
Southern Forest Experiment Station
Starkville, MS 39759-0906

ABSTRACT

Remote sensing represents a useful source of auxiliary information for updating forest inventories. Analysis of remotely sensed imagery provides an efficient means of identifying and quantifying important area changes such as timber harvesting, regeneration, land use conversion, and natural disturbance. The use of remote sensing techniques for updating may involve acquisition of current imagery at the time of an update to improve area estimates, or direct comparison of imagery acquired at two or more points in time to identify specific areas of change and estimate rates of change. It is unlikely that any single remote sensing technique or type of imagery can accomodate all the information requirements of inventory updating.

INTRODUCTION

A basic objective of forest inventory is to provide current information on the status of various forest resources. Successive inventories provide information on the changes or trends in these resources. As changes on forest land occur with increasing frequency, or as time elapses since the last inventory, the need for more current information becomes increasingly critical. In the United States, extensive forest inventories conducted by the U.S. Forest Service are based on remeasurement of scattered permanent sample plots about every 10 years, and remeasurement at shorter intervals is too expensive or impractical. In some areas, natural or human-caused disturbances may cause dramatic changes in the forest land base, causing inventory data to become outdated within a few years. Hence, there is considerable interest in updating techniques that can keep inventory databases current and accurate and that rely on alternative, less-costly sources of information for monitoring forest resources (Birdsey 1989).

Remote sensing represents a useful source of auxiliary information for updating forest inventories. Data collected from airborne and spaceborne sensors can provide timely information on the extent, condition, and dynamics of forested areas. Most forest inventories rely upon remote sensing, usually in the form of aerial photography, for information on forest area (Schreuder and Bonnor 1987). Because remote sensing provides a rapid means of collecting land cover information over large areas, it is often considered an essential component in updating schemes. Important area changes such as timber harvesting, regeneration, land use conversion, and natural disturbances can be detected and measured through appropriate analysis of remotely sensed imagery.

There are many potential applications of remote sensing in inventory updating. The type of remote sensing system required depends upon the level of information desired, the size of the area considered, the cost of acquiring and processing the data, and the ability to perform particular types of analyses (Driscoll et al., 1983, Werth 1983). Techniques range from simple visual interpretation of aerial photography to sophisticated change detection techniques with digital satellite imagery; from detecting subtle changes in stand structure and composition to monitoring broad land cover changes at the county or state level. Remote sensing techniques may be used alone to provide independent estimates of area change, or may be used in conjunction with mathematical models or sample surveys, providing stratification for allocation of field samples.

In any event, there are basically two approaches in the use of remote sensing for updating. First, current imagery may be acquired at the time of an update to provide a current estimate of forest area. Area change is determined by comparing this estimate with that of the previous survey. Specific tracts that have changed during the update period are not identified. Alternatively, imagery acquired at two or more points in time can be registered to each other and compared directly to determine specific areas of change and rates of change. This approach can effectively identify areas that have changed since the last inventory. Because the percentage of area changed in an update period is likely to be relatively small, this approach may be used to efficiently allocate ground sampling primarily to the change stratum.

In this paper, some of the basic tools and techniques of remote sensing are discussed as applied to forest inventory updating. Primary emphasis is on the use of photographic and digital imagery that are commonly available today. Other systems, such as thermal scanners and microwave systems, though important, have either limited application or unproven capabilities in forest inventory and are not discussed here.

PHOTOGRAPHIC SYSTEMS

Manual interpretation of aerial photography and photographic products of satellite imagery is perhaps the most well-known and straightforward method of forest area monitoring by remote sensing. Aerial photography can provide very high ground resolution and can be used for relatively detailed measurements of forest area parameters. Especially when using stereoscopic coverage, the photointerpretation characteristics of shape, size, tone or color, pattern, texture, and site can all be used by the interpreter to arrive at accurate interpretations. When comparing multidate imagery, factors such as atmospheric conditions, sun angle, and vegetation phenology are less of a problem in manual photointerpretation than with digital image analysis (Werth 1983).

Aerial photography has, for many years, been an integral part of most forest inventories for providing forest area information. In the United States, except for Alaska, the basic sampling design for extensive forest surveys consists of double sampling for stratification, in which sample plots on aerial photos are classified into various strata. Permanent plots within these strata are then sampled on the ground. The photointerpretation provides information about land use, timber volume classes, and general forest type categories (Schreuder and Bonnor 1987).

Photointerpretation has been used to advantage in several midcycle update surveys (Scott 1983, Thomas 1983, Thomas and McWilliams 1985). In midcycle surveys of Louisiana (Thomas 1983) and Mississippi (Thomas and McWilliams 1985), high altitude optical bar camera (OBC) panoramic photography was acquired nearly concurrently with the sample surveys and was used to develop updated forest area estimates. Area change was determined by comparison with the previous survey estimates. Also in Mississippi, an intensified sampling of the OBC photography was conducted to provide improved timber removal estimates.

Direct multidate comparison of paired photographic imagery for detecting forest area changes has been less common, perhaps because of difficulty in finding similar coverage (film type, scale, quality, time of year, etc.) with a suitable time interval. A study evaluating various types of imagery for detecting forest disturbances found that most disturbances (79 percent) could be accurately detected and identified on Landsat hardcopy images when compared with aerial photography acquired six years earlier, and that 90 percent of all unidentified disturbances could be detected (Aldrich 1975). The conclusion from this study was that high altitude or satellite imagery could be used to monitor permanent ground sample plots between successive inventories. Timely updates can be accomplished by revisiting sample locations where some kind of disturbance has been detected on the imagery.

An innovative study currently underway in North Carolina involves the use of large, permanently located photoplots to provide information about trends in land cover change (Catts et al., 1987). Detailed land cover categories are photointerpreted and delineated within circular plots on 1:12000 scale color infrared photography. These plots correspond to a ground coverage of 1000 acres and are systematically distributed throughout the State. The first set of photography was acquired in 1985. Acquisition of a second set of photography is planned for 1989. Comparison of photoplots at the two dates will provide a measure of the transitional probabilities for the various cover types. The long-term objective of this project is to integrate routinely available remotely sensed data, including satellite imagery and high altitude photography, to provide timber resource acreage data for updating forest survey statistics.

Another method of using aerial photography for change detection is to digitize the photos using an optical-mechanical scanner and apply change detection algorithms to the resulting digital images. This technique has had limited success because of many confounding factors such as variations in illumination for a given cover type within a photo, variations in exposure between photos, and difficulty in registration due to relief displacement and camera geometry. Also, the problems caused by variations in atmospheric conditions, sun angle, and phenology, mentioned earlier, arise here as well. However, this technique has been shown to have some value for detecting significant disturbances such as clearcuts and land clearings (LePoutre 1981).

DIGITAL SYSTEMS

Digital image data are well suited for automated change detection techniques. Although most applications of digital imagery for change detection have used data from the Landsat Multispectral Scanner (MSS) or Thematic Mapper (TM) sensors, data from other sensor systems, such as SPOT or the Advanced Very High Resolution Radiometer (AVHRR), could be used as well. Conceptually straightforward, digital change detection procedures are based on the premise that a change in the spectral response of a pixel over time can be related to a corresponding change in land cover. Theoretically, if two images can be precisely registered to each other, a pixel-by-pixel comparison (done by a computer) should reveal areas where some kind of change has occurred. In reality, however, the difference in spectral radiance measured by a pixel can be affected by a number of factors not related to land cover change, such as atmospheric attenuation, vegetative phenology, soil moisture, sun angle, and changes in sensor geometry and calibration (Burns 1983, Driscoll et al., 1983). Also, errors in registration may produce some false indications of change. Some of these factors may be controlled by choosing a sensor system that records imagery of the same geographic area at the same time of day with consistent geometric and radiometric characteristics, and by selecting particular anniversary dates to minimize the effects of phenology and soil moisture (Jensen 1986). Still, digital change detection can be a complicated process involving consideration of many details.

Various digital change detection procedures have been developed, ranging from simple to complex. Three of the more commonly used procedures are reviewed below:

Image Differencing and Ratioing

Image differencing involves subtracting the digital value of a pixel in a given spectral band at time 2 from the corresponding pixel value at time 1. Large resulting positive or negative values indicate change, whereas the expected value of unchanged pixels is zero. A constant is usually added to the output data set to avoid negative values. Similarly, image ratioing consists of dividing the value of a pixel at time 2 by the corresponding value at time 1. Unchanged pixels have an expected value of 1. The output data are usually multiplied by a constant to expand the range of values. A variation of this approach, difference of ratios, involves calculating the ratio of two different bands (such as infrared/red) for a pixel at time 1 and time 2 and taking the difference. Image differencing and ratioing are relatively easy to implement and are useful for quickly detecting changed areas, but do not provide information on the nature of the change or the resulting land covers. They have been used successfully for detecting severe insect defoliation over a one-year period using Landsat MSS data (Nelson 1982).

Delta Classification

This procedure involves comparing two independently classified data sets. Change is indicated by a shift in categorical value of a pixel between time 1 and time 2. The procedure has the advantage of identifying areas that have changed as well as the type of change that has occurred. A drawback is that very accurate classifications are required for both data sets, or else changes will be indicated falsely. High classification accuracy cannot always be attained and even with moderately accurate classifications, errors are likely to be compounded in the output data set. This has been found to lead to a general overestimation of change (Colwell et al., 1980). Despite its problems, this procedure is commonly used because of its simplicity and potential information content. A number of variations have been suggested to alleviate its inherent problems (Colwell et al., 1980, Nelson 1982).

Change Vector Analysis

Change vector analysis is similar to image differencing in that the magnitude of the difference in value between two points in time is used to indicate whether or not a change has occurred. In this case, however, the direction of the change in spectral space is used to indicate the type of change that has happened (Colwell et al., 1980, Malila 1980). This procedure is usually conducted in several steps. First, the original number of spectral bands (at least four, when using Landsat) from each date are collapsed into two transformed spectral variables (brightness and greenness). This reduces the dimensionality of the problem but preserves most of the information about vegetation (Kauth and Thomas 1976). Next, a spatial averaging algorithm is applied to the data to identify clusters of pixels that are spectrally uniform and spatially contiguous on both data sets (Kauth et al., 1977). Average brightness and greenness values are then computed for each date for each unit generated by the spatial averaging algorithm. This also reduces the size of the analysis by shifting from a per-pixel basis to a spatial unit (cluster of pixels) basis. Finally, vectors describing the magnitude and direction of change from time 1 to time 2 are computed for each spatial unit. By comparing the magnitude of change with a specified threshold, a decision can be made as to whether a unit has changed. If a unit has changed, the direction component of the change vector can be examined to derive information on the type of change. For example, an increase in greenness and decrease in brightness corresponds to increasing vegetation coverage. Conversely, a decrease in greenness and increase in brightness can indicate deforestation. Although this technique appears to have good potential for monitoring forested areas, it has not been widely adopted, perhaps because of its complexity. There are also certain shortcomings, such as an inability to distinguish real land cover changes from transient, within-type spectral variations. This procedure has been demonstrated with reasonable success on study areas in South Carolina (Colwell et al., 1980) and Idaho (Malila 1980).

Each of the above procedures has its own particular advantages and disadvantages. All of them, however, are capable of indicating where and how rapidly major changes in forest area are occurring. Digital change detection procedures have several important features from the standpoint of updating forest inventories, including:

- The repetitive, large-area coverage of satellite data. This represents a continuing source of land cover information over extensive areas that can be used for updating purposes;

- The use of computer-aided analysis techniques, allowing for rapid processing of large amounts of data and directing attention to areas indicating change; and

- The ability to regularly update land cover layers in a geographic information system - easily accomplished because the data are already in digital form.

CONCLUSIONS

Remote sensing presents a wide variety of tools and techniques that can contribute to forest inventory updating. Each remote sensing system has its own strengths and weaknesses and no single type of remotely sensed data or analysis technique is likely to be optimal in all updating situations (Driscoll et al., 1983). Probably the greatest potential for large-scale inventory updating lies in the use of digital change detection techniques. However, much work remains to identify optimal change detection algorithms for specific geographic areas and problems (Jensen 1986). Also, a computer lacks the reasoning ability of a skilled photointerpreter, and the judgment and experience of an analyst is still tremendously important even when using computerized methods.

An efficient updating system is likely to involve more than one level of remotely sensed data, each level providing the type of information for which it is best suited (Schreuder et al., 1986). Satellite imagery can provide change information at low spatial resolution but high temporal resolution for the entire population, whereas aerial imagery can provide more detailed information on a subsample, all of which will ultimately be linked to ground information. This information may be used to provide strata for sample survey updates, or may be input directly to modelling efforts by providing exogenous estimates of harvest rates, regeneration rates, and other significant trends that span the intersurvey years.

In addition to the need to monitor forests for local and national planning purposes, remote sensing may also help in refining biome-level estimates of deforestation. Satellite data from the coarse-resolution AVHRR sensor is particularly well suited for this application (Malingreau and Tucker 1987). This type of information may help answer questions about global changes in the amount of carbon held in terrestrial ecosystems.

LITERATURE CITED

Aldrich, R.C., 1975: Detecting disturbances in a forest environment. Photogrammetric Engineering and Remote Sensing, vol. 41, no. 1, pp. 39-48.

Birdsey, R.A., 1989: A dynamic matrix model for updating forest resource inventories. PhD Dissertation, State University of New York, College of Environmental Science and Forestry, Syracuse, NY, 267 pp.

Burns, G.S., 1983: Land cover change monitoring within the east central Louisiana study site - a case for large area surveys with Landsat Multispectral Scanner data. Technical Report No. DC-Y3-04418-NSTL/ERL-221. NSTL Station, MS: National Aeronautics and Space Administration, National Space Technology Laboratories, Earth Resources Laboratory, 31 pp.

Catts, G.P., Cost, N.D., Czaplewski, R.L., and Snook, P.W., 1987: Preliminary results from a method to update timber resource statistics in North Carolina. In: Color Aerial Photography and Videography in the Plant Sciences and Related Fields: proceedings of the Eleventh Biennial Workshop on Color Aerial Photography in the Plant Sciences; 1987 April 27-May 1; Weslaco, Texas. Falls Church, VA: American Society for Photogrammetry and Remote Sensing, pp. 39-52.

Colwell, J., Davis, G., and Thomson, F., 1980: Detection and measurement of changes in the production and quality of renewable resources. Final Report, Report No. 145300-4-F, Contract No. 53-3187-9-47. Environmental Research Institute of Michigan, Ann Arbor, Michigan. Houston, TX: U.S. Department of Agriculture, Forest Service, Nationwide Forestry Applications Program, 144 pp.

Driscoll, R.S., Frayer, W.E., and Werth, L.F., 1983: Complexities of remote sensing with sampling for change detection. In: Proceedings, Renewable Resource Inventories for Monitoring Changes and Trends; 1983 August 15-19; Corvallis, Oregon. Corvallis, OR.: Oregon State University, pp. 613-617.

Jensen, J.R., 1986: Introductory Digital Image Processing. Prentice-Hall, Englewood Cliffs, NJ, 379 pp.

Kauth, R.J., Pentland, A.P., and Thomas, G.S., 1977: BLOB: An unsupervised clustering approach to spatial processing of MSS imagery. In: Proceedings, Eleventh International Symposium on Remote Sensing of Environment; 1977 April 25-29; Ann Arbor, Michigan. Ann Arbor, MI: Environmental Research Institute of Michigan, pp. 1309-1317.

Kauth, R.J., and Thomas, G.S., 1976: The Tasseled-Cap -- a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In: Proceedings, Machine Processing of Remotely Sensed Data; 1976 June 29-July 1; West Lafayette, Indiana. W. Lafayette, IN: Purdue University, Laboratory for Applications in Remote Sensing, pp. 4B41-4B51.

LePoutre, D.L., 1981: Monitoring disturbances and change on the San Juan National Forest. Final Report, Report No. NFAP-250, Contract No. 53-3187-1-42; Houston, TX; U.S. Department of Agriculture, Forest Service, Nationwide Forestry Applications Program, 82 pp.

Malila, W.A., 1980: Change vector analysis: an approach for detecting forest changes with Landsat. In: Proceedings, Machine Processing of Remotely Sensed Data; 1980 June 3-6; West Lafayette, Indiana. W. Lafayette, IN: Purdue University, Laboratory for Applications in Remote Sensing, pp. 326-335.

Malingreau, J.P., and Tucker, C.J., 1987: The contribution of AVHRR data for measuring and understanding global processes: large-scale deforestation in the Amazon Basin. In: Proceedings, International Geoscience and Remote Sensing Symposium - IGARSS '87; 1987 May 18-21; Ann Arbor, Michigan. New York, NY: Institute of Electrical and Electronic Engineers, pp. 443-448.

Nelson, R.F., 1982: Detecting forest canopy change using Landsat. Final Report, Report No. NFAP-301, Contract No. RR-G2-04258. Greenbelt, MD: National Aeronautics and Space Administration, Goddard Space Flight Center, Earth Survey Applications Division, Earth Resources Branch, 80 pp.

Schreuder, H.T., Snook, P.W., Czaplewski, R.L., and Catts, G.P., 1986: A proposed periodic national inventory of land use land cover change. In: Proceedings, 1986 American Society for Photogrammetry and Remote Sensing -American Congress of Surveying and Mapping Fall Convention, ASPRS Technical Papers; 1986 September 28-October 3; Anchorage, Alaska. Falls Church, VA: American Society for Photogrammetry and Remote Sensing, pp. 255-263.

Schreuder, H.T. and Bonnor, G.M., 1987: Forest inventories in the United States and Canada. Forestry Chronicle, vol. 63, no. 6, pp. 431-434.

Scott, C.T., 1983: Example of midcycle updating to assess catastrophic change. In: Proceedings, Renewable Resource Inventories for Monitoring Changes and Trends; 1983 August 15-19; Corvallis, Oregon. Corvallis, OR: Oregon State University, pp. 555-557.

Thomas, C.E., 1983: Maintaining midcycle inventories of forest resources. In: Proceedings, Renewable Resource Inventories for Monitoring Changes and Trends; 1983 August 15-19; Corvallis, Oregon. Corvallis, OR: Oregon State University, pp. 386-390.

Thomas, C.E. and McWilliams, W.H., 1985: Midcycle survey of Mississippi's forest resources. Resource Bulletin SO-101. New Orleans, Louisiana: U.S. Department of Agriculture, Forest Service, 37 pp.

Werth, L.F., 1983: Remote sensing for forest land change detection--an approach for national assessment. In: Proceedings, Renewable Resource Inventories for Monitoring Changes and Trends; 1983 August 15-19; Corvallis, Oregon. Corvallis, OR: Oregon State University, pp. 330-333.

Inventory Updates: Overview of The North Central Experience

W. Brad Smith
Research Forester
USDA, Forest Service
North Central Forest Experiment Station
St. Paul, Minnesota 55108

ABSTRACT

Forest growth models are used by Forest Inventory and Analysis (FIA) units to estimate current conditions by updating past forest inventories. Applications range from integrating growth models with field inventories to providing multi-state common year inventory data for the RPA Assessment. Proper accounting of removals, growth, mortality, and land use change is critical in an update process. This paper presents two methods used by the North Central Experiment Station to address these components of change in updating forest inventories with the STEMS growth model.

INTRODUCTION

Forest growth modelling can be applied in many ways, from projecting silvicultural responses of individual stands to mid-cycle updating of state inventories to predicting future timber supplies for entire countries. This paper will look at the basic elements of one application -- updating forest inventories and how they are addressed by the Forest Inventory and Analysis (FIA) staff at the North Central Forest Experiment Station.

The term "update" is defined as an estimate of the current forest condition derived by modelling the dynamic change in a forest from a known time and condition in the past. The major components of inventory change are:

REMOVALS	(from continuous forest land)
GROWTH	(survivor growth, ingrowth, growth on removals and mortality)
MORTALITY	(normal and catastrophic losses)
LAND USE CHANGE	(landclearing and reforestation)

The ultimate update process would predict exactly when and where each of these actions occurred and would report the consequences. In updating, however, we generally settle for a process that provides a fair rendition of the information that would be available from a new field inventory.

BACKGROUND

If our goal is to produce a simulated field inventory, we must first review what is produced by a full field inventory of the forest resource that we wish to reproduce. All Forest Service FIA units, except Alaska, use a two-phase sample design (Cochran 1977) consisting of a remote image phase for stratification and a ground phase for verification. In the remote phase (aerial photos or satellite imagery), a systematic random sample of points is selected from the imagery, and land use information is recorded for each point. The remote image phase establishes the preliminary area breakdown of the land base differentiated by forest and nonforest land uses. During the ground phase, a subset of the remote image points is visited to collect data describing the condition of the land and vegetation. Information from the ground phase is used to adjust the remote phase area estimate for misclassifications and for changes that have taken place since the date of the imagery. Other information collected in the ground phase for forested plots includes the species, size, quality, composition, and productivity of the vegetation found. The species and size of harvested and dead trees are also recorded. Information for other local studies related to recreation, wildlife, ecology, and other disciplines may also be collected.

The time and expense of full field inventories can be considerable. The cost for a standard FIA field inventory can easily exceed $90 per thousand acres of forest land inventoried, and fieldwork for a State or region with several million acres of forest land can consume years. The cost of a large-scale computer update, on the other hand, can generally be held to under $5 per thousand acres of forest land including the cost of model calibration and validation. Supplemental fieldwork in conjunction with a computer update could add another $5 per thousand acres of forest land. Before you ditch all those expensive field inventory plots, remember that reliable modelling requires a data base of reasonably current remeasurement data to calibrate and validate the model for local conditions. Why bother modelling if you still have to inventory? Growth models can provide resource analysts and planners with valuable mid-cycle inventory data at low cost. A second cost saving benefit comes from integrating a growth model into the inventory design so that the number of plots measured and thus direct field costs could be reduced by as much as 20 percent.

Dozens of computer growth models are available. How do you select the most appropriate one? The advent of electronic data processing spawned a generation of sophisticated computer-based models to address inventory change. One of the earliest computer models used by the Forest Service was the Timber Resource Analysis System, TRAS (Larson and Goforth 1970). This model, based on concepts developed by Meyer (1952), is a stand projection model that manipulates removals, mortality, and net growth rates by 2-inch diameter classes for softwoods and hardwoods. TRAS has been the workhorse of Forest Service modelling for more than 20 years, but it is too aggregate for many current applications. Most other early models like Buckman's red pine model (1962) and Clutter's loblolly pine model (1963) operated at the stand level. These models were an improvement but still too confining for stands with highly mixed species and size structure. Individual tree models such as PROGNOSIS (Stage 1973) and STEMS (Belcher et al. 1982, Shifley 1987), which could handle a wide variety of stand conditions and provide updated tree lists as outputs, emerged in the mid-1970's. The ability to produce updated tree lists made these latter models an ideal choice for FIA applications because updated and remeasured data could be summarized in the same manner. The choice of model is mandated by the level of detail required in the model output.

UPDATING AT NORTH CENTRAL

The North Central FIA uses two methods for updating inventories with STEMS. Short-term updates incorporate exogenous data to account for removals and land change; model-integrated inventories use specialized sampling techniques to assess change. The following sections will review the four major components of inventory change and the way they are handled by North Central FIA for short-term updates and the STEMS-integrated inventory design.

Removals

We often think of modelling as nothing more than predicting growth and mortality. In a way that is true, but the first thing we ask for when applying the model is a way to cut trees during projections or account for trees that have been cut in the case of updates. Because removals will generally range from 40 to 100 percent of growth, removals estimation literally becomes a modelling problem in itself. Thus, although it may seem odd to discuss removals first, it is the biggest problem in updating, short of having a suitable model.

Two methods have been used at North Central to account for removals during a modelled inventory update, depending on time, money, and the availability of information. A "cognitive" method is used with our STEMS-integrated inventory design (Hansen and Hahn 1983), and a "blind" method is used for short-interval common year updating such as that done for the RPA Assessment (USDA 1988, Smith and Hahn 1986). A brief description of each approach follows:

Cognitive: The STEMS-integrated inventory design is a variant of our standard two-phase sample (Fig. 1) which stratifies the ground-check photo plots from the previous inventory as to whether or not they are "disturbed." The term "disturbed" refers to any condition of change since the last inventory that can not be simulated by the growth model. This would include plots that were seedling stands at the last inventory, stands that have been clearcut or detectably thinned, or stands at a high risk of change (high probability of catastrophic loss between time of remote imagery and the expected time of ground verification). An example of high risk would be an area where spruce budworms may have become epidemic since remote imagery was collected. In this case ground plots in balsam fir stands would be classed as "disturbed" even though there is no visible evidence of disturbance on the remote imagery. Removals caused by catastrophic loss are accounted for in this manner. A weak link in this process is the case of "dispersed catastrophes," which occurs when there is volume loss at epidemic levels that is not obvious when viewing the remote imagery. Dutch elm disease is a clear example of this problem. Elm occurs in highly mixed stands in the East. Even when mortality is above 50 percent in the species, it appears to be nothing more than random mortality in the stand and the plot is not classified as disturbed. When conditions suggest a dispersed catastrophe, plots in highly susceptible categories should be classified as disturbed without regard to their appearance on the remote imagery. This determination is based on a review of data from the previous inventory.

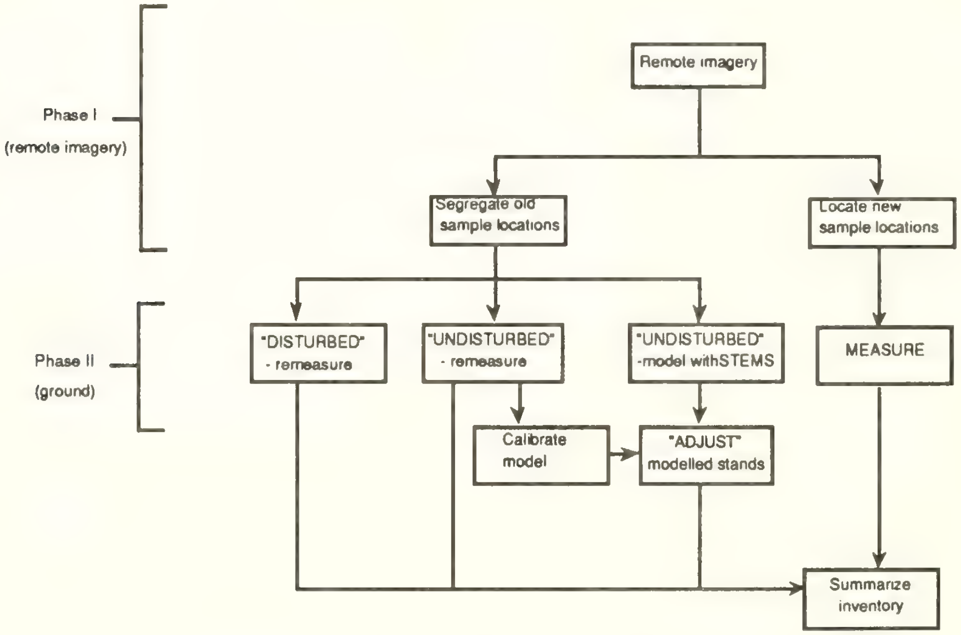


Figure 1.- Simplified STEMS-integrated inventory design

Blind: Unlike the cognitive method, the blind update method (Fig. 2.) has no remote imagery phase and estimates of removals and land change are derived from external sources (Smith and Blyth 1989). FIA at North Central has devoted much effort to developing methods of monitoring timber removals on an annual basis through annual production studies for pulpwood , periodic mill production studies, and periodic fuelwood production studies. Data from these studies are linked to timber utilization studies (Blyth and Smith 1980) to synthesize an accurate picture of annual removals by species and Forest Survey Unit. When possible we segregate removals by ownership class to improve allocation during the update.

Computerized management guides (Brand 1981, Walters 1988) are then used to select a subset of inventory plots that are eligible for silvicultural treatment during the update period. These guides are used in lieu of detailed information on actual management strategies. The initial plot list is sorted in descending order by age, presuming that oldest stands will be harvested first. A special volume control removals algorithm scans each selected plot to determine if it will be cut. The volume from plots selected for cutting are accumulated until the target volume of removals by species and FIA survey unit is reached. Once the target is reached, no further plots are scanned for harvest opportunity.

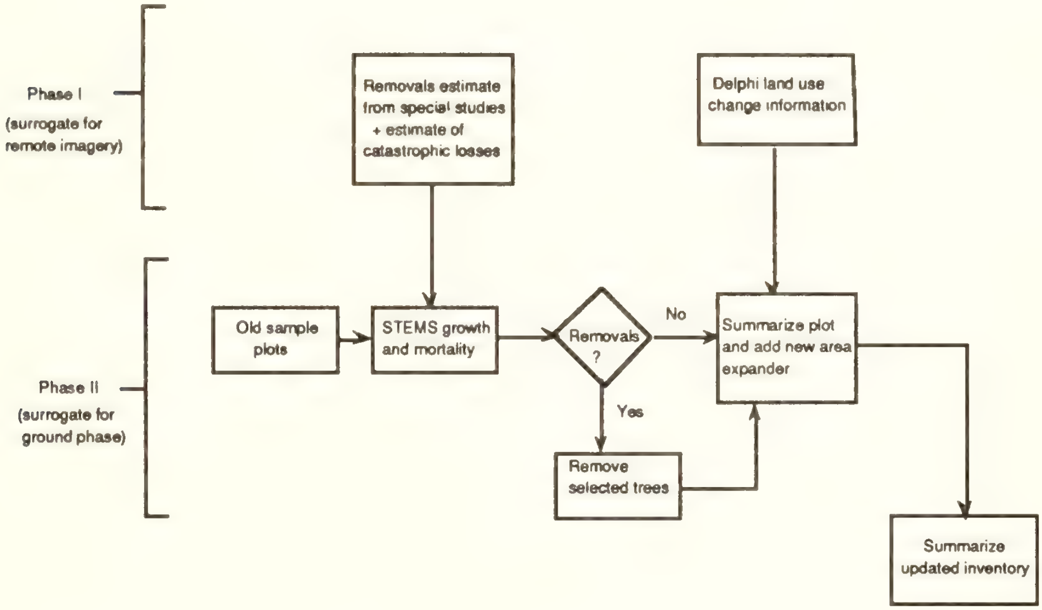

Phase I
(surrogate for remote imagery)
Removals estimate from special studies + estimate of catastrophic losses
Delphi land use change information
Phase II
(surrogate for ground phase)
Old sample plots
STEMS growth and mortality
Removals ?
No
Yes
Summarize plot and add new area expander
Remove selected trees
Summarize updated inventory

Figure 2.- Simplified overview of the 'blind' update process

The blind update is clearly a rough-cut approach with limited geographic reference to the location of actual removals. This method might be improved by establishing studies to look at the characteristics of remeasurement plots with evidence of cutting and relate this information to simulated prescriptions selected by a computer management algorithm. The resulting relationships would then be used to develop a probability of harvest function that could be used to more realistically allocate removals for blind updates. The Pacific Northwest FIA unit has probably done more work in this area than any other FIA unit.(MacLean and McKay 1989). Another alternative developed by the FIA unit at the Southern Station (Thomas 1983) uses a sub-sample of remeasurement plots to develop regression estimators for current volume and growth per acre. This information, combined with land change estimates from remote imagery, generates an estimate of net change and implicitly accounts for removals and mortality.

Growth

The growth functions for STEMS were calibrated and validated with data from throughout the Lake and Central States (Holdaway and Brand 1986, Shifley 1987). Test projections over a range of forest conditions produced reliable results when compared with remeasurement data from each region.

In all updates, recent remeasurement data are used (when available) to calibrate and validate the model for local conditions. Three methods have been developed at North Central to address local growth calibration. The first method is a simple ratio-of-means estimator that is applied as a multiplicative adjustment by species and broad diameter classes to the annual predicted increment (Smith 1983). The second method, proposed by Holdaway (1985), is a continuous function that generates an additive increment adjustment as a function of species and dbh. The third method, developed by Hansen and Hahn (1983) in conjunction with the STEMS-integrated inventory design, applies an adjustment to stand volume by forest type based upon regression estimators derived from a subset of "undisturbed" remeasurement plots and proportionally allocates the adjustment to each tree to provide updated tree lists. Of the three methods, ratio-of-means is used most frequently because of its simplicity and effectiveness for short-range modelling applications.

Although STEMS regeneration routines are not fully developed at this time, the outcome of short updates for timber volume estimation is generally not significantly affected. Most trees that were established on harvested land or in seedling stands at the beginning of the update will not grow to merchantable size by the end of the period. A regeneration model becomes critical when the desired projection period exceeds the time necessary for significant numbers of seedlings to reach the minimum merchantable size.

Mortality

STEMS mortality functions have been developed for most major species in the North Central region (Buchman et al. 1983, Buchman and Lentz 1984). These functions are intended to predict normal random mortality including the impact of endemic pests, pathogens, and climatic stresses. Catastrophic mortality resulting from fires, windthrow, epidemic insect outbreaks, or diseases must be handled as unique events in the update process. The STEMS-integrated inventory design accounts for most of these events through disturbance detection in the remote image sampling phase. Dispersed catastrophes must be addressed by declaring old plots to be disturbed as previously described. Blind updates would use a Delphi approach (Gregersen et al. 1989) to arrive at expert opinion of catastrophic losses and then adjust removals targets for those species to achieve a "proper" loss rate. Ratio-of-means estimators are derived from remeasured data to calibrate mortality functions for local conditions.

Land Base Change

Land base change is handled directly in the STEMS-integrated inventory design during the remote image sample phase. Old plots whose land use patterns have changed are classified as disturbed and are remeasured in the field. New photo work for a short-term blind update would be too costly and time consuming. We've found that for update periods of less than 10 years, area change can be estimated using trend analysis from previous inventory data and gathering expert opinion from local resource managers. These data then provide surrogate remote-phase area estimates by ownership and forest type to be used with the updated plot data. The objectives of the diverse group of owners that hold forest land play a significant role in the stability of the timberland base. It is always wise when releasing updated inventory data to caution readers to contact the appropriate agency to verify figures for any analysis that may be sensitive to local changes in the timberland base, such as recently designated reserved or deferred forest land. Area and volume figures would have to be adjusted accordingly for any parcels of land affected by significant changes in the timberland base. Additional information about non-industrial private owners and their objectives may also be available from forest land ownership studies.

APPLICATIONS

This paper briefly summarized two updating procedures used by FIA at North Central. To date we have used the STEMS-integrated inventory design in the 1981 Wisconsin FIA inventory as well as in the current Missouri and Minnesota inventories in progress. To prepare for the 1990 RPA Assessment we have conducted blind updates on 9 of the 11 state inventories in our region. These updates brought the most recent inventory data for each state up to the 1987 base year for the Assessment.

Novices to the world of modelling and updating should gather as much background information about their application as possible. Once you have selected a potential model it would be useful to read Buchman and Shifley (1983) and Brand and Holdaway (1983). These readings will help you determine what to expect from the model in terms of reliable growth and mortality. As a rule, most growth models are designed to treat individual plots or stands and will not provide mechanisms to handle timber removals or land use change. STEMS has computerized timber management algorithms, but they are designed to suggest sound silvicultural practices rather than mimic actual historic practices. The technology of updating inventories will continue to evolve, and the focus of new research will shift with each improvement in our ability to address any of the four major components of change.

REFERENCES

Belcher, David W.; Holdaway, Margaret, R.; Brand, Gary J. 1982. A description of STEMS: The stand and tree evaluation and modelling system. Gen. Tech. Rep. NC-79 St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station. 18 p.

Blyth, James E.; Smith, W. Brad. 1980. Minnesota logging utilization factors, 1975-1976: development, use, implications. Res. Bull. NC-48. St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station. 8 p.

Brand, Gary, J. 1981. Simulating timber management in Lake States forests. Gen. Tech. Bull. NC-69 St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station. 25 p.

Brand, Gary J.; Holdaway, Margaret R. 1981. Users need performance information to evaluate models. Journal of Forestry. 81(4):232-234.

Buchman, Roland G.; Pederson S. P.; Walters, Nancy R. 1983. A tree survival model with application to species of the Great Lakes Region. Canadian Journal of Forest Research 13(4):601-608.

Buchman, Roland G.; Shifley, Stephen R. 1983. Guide to evaluating forest growth projection systems. Journal of Forestry 81(4):pp232-234.

Buchman, Roland G.; Lentz, Ellen L. 1984. More Lake States survival predictions. Res. Note NC-312. St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station. 6 p.

Buckman, R. E. 1962. Growth and yield of red pine in Minnesota. USDA Tech. Rep 1272. St. Paul, MN: U.S. Department of Agriculture Forest Service, Lake States Forest Experiment Station. 50 p.

Clutter, Jerome. L. 1963. Compatible growth and yield models for loblolly pine. Forest Science. 9(3):pp355-371.

Cochran, William G. 1977. Sampling techniques. New York, NY: John Wiley & Sons, Inc. 428p.

Gregersen, Hans; Lundgren, Allen; Jakes, Pamela. 1989. Identifying emerging issues in forestry as a tool for research planning: A case study of National Forest management and use. Gen. Tech. Rep. NC-137 St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station. (In prep)

Hansen, Mark H.; Hahn, Jerold T. 1983. Estimation of sampling error associated with timber change projection simulators. In Renewable resources inventories for monitoring changes and trends. Ed. John F. Bell and Toby Atterbury. College of Forestry, Oregon State University, Corvallis, Oregon. pp546-549.

Holdaway, Margaret R. 1985. Adjusting STEMS growth model for Wisconsin forests. Res. Pap. NC-267 St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station. 8 p.

Holdaway, Margaret, R.; Brand, Gary J. 1986. An evaluation of the Lake States STEMS85. Res. Pap. NC-269 St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station. 10 p.

Larson, Robert W.; Goforth, Marcus H. 1970. TRAS: A computer program for the projection of timber volume. Agric. Handb. 377. U.S. Department of Agriculture, Forest Service, Washington D.C. 24p.

MacLean, Colin; McKay, Neil. 1989. Updating inventories in western Oregon and Washington using tree growth, mortality, and harvest models developed from permanenent plot data. In these proceedings.

Meyer, A. H. 1952. Structure, growth, and drain in balanced uneven-aged forests. Journal of Forestry 50:85-92.

Shifley, Stephen R. 1987. A generalized system of models forecasting Central States tree growth. Res. Pap. NC-279 St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station. 10 p.

Smith, W. Brad. 1983. Adjusting the STEMS regional forest growth model to improve local predictions. Res Note. NC-297 St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station. 5 p.

Smith, W. Brad; Blyth, James E. 1989. Timber harvesting trends in the Lake States, 1983-1987. Res Pap. NC-288 St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station. 18 p.

Smith W. Brad; Hahn, Jerold T. 1986. Michigan's forest statistics, 1987: An inventory update. Gen. Tech. Rep. NC-112. St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station. 44 p.

Stage, Albert R. 1973. Prognosis model for stand development. Research Paper INT-137. Ogden, UT: U.S. Department of Agriculture Forest Service, Intermountain Forest and Range Experiment Station. 33 p.

Thomas, Charles E. 1983. Maintaining midcycle inventories of forest resources. In Renewable resources inventories for monitoring changes and trends. Ed. John F. Bell and Toby Atterbury. College of Forestry, Oregon State University, Corvallis, Oregon. pp. 386-390.

USDA Forest Service. 1988. An analysis of the timber situation in the United States: 1989-2040. Part I: The current resource and use situation. Forest Service Report (DRAFT), Washington D.C.; U.S. Department of Agriculture, Forest Service. 529p.

Walters, Nancy R. 1988. Central States forest management guides as applied in STEMS. Gen. Tech. Rep. NC-119 St. Paul, MN: U.S. Department of Agriculture Forest Service, North Central Forest Experiment Station. 18 p.

Updating inventories in western Oregon and Washington using tree growth, mortality and harvest models developed from permanent plot data.

Colin D. MacLean
Research Forester
Pacific NW Research Station, U.S. Forest Service
Portland, OR 97208

Neil McKay
Research Forester
Pacific NW Research Station, U.S. Forest Service
Portland, OR 97208

ABSTRACT

Researchers at the USDA-Forest Service Forest Inventory and Analysis Unit (FIA) in Portland, OR are developing methods for updating FIA inventories in western Oregon and Washington. Techniques are described for modeling diameter and height growth, mortality, and harvest effects. The models are developed using remeasured data from the inventories that require updating. The models provide estimates of growth, mortality, and harvest that are used to update plot and tally tree data. Updates are produced using the same computer programs used to compile the original inventories. The techniques will enable the Portland FIA unit to update inventories that are up to 10 years old.

INTRODUCTION

When inventories no longer reflect the current state of the forest resources, their useful live can be extended through updating. By updating, information needs often can be met at a substantial cost savings over remeasurement. In addition, updating can enhance the quality of regional and national assessments since inventories varying in age can be made current to a specified base year.

The USDA--Forest Service Forest Inventory and Analysis Unit (FIA) in Portland, OR conducts inventories in the Douglas-Fir subregion--western Oregon and Washington. Currently, these inventories cover all lands except National Forests and lands in western Oregon managed by the USDI--Bureau of Land Management. The inventories are based, in part, on a grid of permanent plots. Field crews visit the plots periodically, recording changes in tree size, losses of tally trees to mortality and harvest, and changes in land use. Since visits are approximately ten years apart, inventories, at any one time, vary from current to ten years old. Shortening the time span between visits is not feasible due to technical, budgetary and administrative constraints.

Within the Douglas-fir subregion, trees grow rapidly and harvest activities are intense. Thus, FIA inventories in the subregion become obsolete before the ten-year measurement period has elapsed. For this reason, the FIA staff in Portland places high priority on developing techniques to provide mid-cycle updates of its most recent inventories in the subregion.

THE PROBLEM OF UPDATING

Aggregated inventory statistics are inadequate to supply the information needs of FIA's diverse users. Sophisticated users such as researchers and consultants often require complete plot and tree files. Thus, an update should include revision of all plot and tree variables that might have changed since the last measurement. At FIA-Portland, we are developing statistical models that allow us to update tree and plot variables for every plot in the original inventory. Once the variables are updated, we can compile an updated inventory using the same computer programs and providing the same outputs as the original inventory.

Smith, of the FIA unit in St. Paul, Minnesota, describes an update system that meets these updating needs through use of STEMS--a comprehensive, individual tree simulation model capable of long-term projections as well as short-term updates (Smith, these proceedings). Shifley and Fairweather (1983) developed tree growth and mortality models for a Douglas-fir subregion version of STEMS, a variant that has had only limited testing, lacks a harvest model, and is calibrated regionally rather than by specific inventory unit.

Other models are designed primarily for long-term projections or lack the levels of resolution required for updating. DFSIM is a stand model calibrated only for young, managed stands of Douglas-fir (Curtis et. al. 1981). Two individual tree-based models--SPS (Arney 1985) and Organon (Hester et. al. 1988)--provide the user with updated tree lists but are calibrated for ranges of stand conditions and local conditions that are narrower than those sampled by FIA inventories.

Because these models are available, developing new long-term projection simulators is not a high priority research need of the Portland FIA unit. Instead, we are concentrating a simpler objective: how to update our most recent inventories for periods 10 years or less using models calibrated to our most recent successive measurements data.

THE FIA-PORTLAND APPROACH

First, we develop estimators for tree growth, tree mortality and tree harvesting from data gathered during two successive measurements of the inventory sample. These estimators are used to estimate, by extrapolation, tree growth, mortality, and harvest during the update period. When compiling an inventory, we develop models to estimate the current rate of change in diameter and height growth and mortality of each live tally tree. These same models will later provide the means of updating these trees for change that occurs after the inventory. The growth equations need no additional calibration since they are, for the most part, applied to the same trees that were used to develop the equations.

We account for harvest effects by estimating, for each timberland plot, the probability that the plot will be harvested during the update period. These probabilities are used to allocate independent estimates of aggregate volume harvested during the period. The estimators for harvest probabilities are developed from harvest data collected during the two most recent inventories.

Two other components of change--land use shifts and regeneration success--follow historical trends and do not require formal models.

The FIA methods for updating are tailored to the conditions that occur in the Douglas-fir subregion. Land use changes are predictably minor. On private lands in western Oregon, for example, only about 12,000 acres of forest land have been converted to nonforest land in each of the past 25 years (MacLean, in press). The prevalent method of harvest -- clearcutting -- permits modeling at plot-level rather than the tree-level; we can ignore the presence of any residual trees left after harvest. The impacts of partial harvesting on forest conditions are so minor that they can be omitted without appreciable error. Because stands are predominantly even-aged, regeneration is important only when existing stands are removed by clearcut harvest or natural catastrophe. Since Oregon and Washington require the planting of conifers where conifer stands were harvested and since inventory examinations confirm that harvested acreage is invariably replanted, we can assume that adequate regeneration follows wherever harvest is estimated to occur.

METHODS

To develop an update system for FIA inventories, we are constructing four types of models: (1) a DBH growth model, (2) a tree height growth model, (3) a tree mortality model, and (4) a model of harvest effects. Regeneration and land use change estimates are based on broad assumptions. A description of these components follows:

Regeneration assumptions

We assume: (1) that all harvesting is done by clearcutting; (2) that all trees are harvested or killed; (3) that the acreage cut is regenerated within 5 years following harvest; (4) that all trees in these regenerated clearcuts remain seedlings--any possible update period is too short to grow trees to a larger size; (5) that the species planted depends on the site and local planting practice but is almost always a conifer; (6) that, because of the even-aged nature of most stands in the Douglas-fir subregion, volunteer seedlings in established stands seldom have an important effect on stand characteristics; and (7) that nonstocked areas other than recent clearcuts remain nonstocked during an update period--a reasonable assumption given the short time span of our updates and the heavy inhibiting vegetation usually found on older nonstocked areas.

Land use change

Analysis of successive FIA inventories in the subregion indicates that land use change occurs almost entirely in one direction--forest to nonforest--at a low and predictable rate. Were we to assume that forest land area remains constant, we would overestimate the area when updating. For western Oregon, the overestimation would be about 0.2 percent for each year of an update. We could correct for this bias by developing a model to estimate the probability of land use change. However, because the data base available for developing a model is very small, correlations between land use change and specific land and forest characteristics would be difficult to discern. At present, we are reducing total forest land area by a constant rate that is based on the historic rate of land conversion. Were this rate more variable, or in both directions--nonforest to forest as well as forest to nonforest--or more correlated with economic conditions, a more sophisticated technique would be necessary.

DBH growth

We assume that, for any given tally tree, basal area growth is constant. Thus, if ADD^2 = Annual change in $(D)^2$, then:

$$ADD^2 \quad = \quad [(D_2)^2 - (D_1)^2]/Y_c$$

and the updated DBH can then be calculated by:

$$D_u \quad = \quad SQRT[(D_2)^2 + (ADD^2)(Y_u)],$$

where: D_1, D_2, D_u = DBH at the first and second measurement, and update year

Y_c, Y_u = Years between measurements and since last measurement.

The assumption of constant basal area (squared diameter) growth may be unwarrented. A recent study (MacLean and Scott 1988) of remeasurement data from western Oregon indicates that basal area growth increased slightly in conifers as they grew larger and declined in hardwoods. The findings are based on three measurements over a 23-year period and may reflect cyclical variation in tree growth. After the next western Oregon inventory, we will again investigate these growth assumptions. In the meantime, the assumption of constant basal area seems a reasonable approximation since an altered growth assumption based on the Oregon findings would seldom affect updated DBH estimates by more than 0.1 or 0.2 centimeters in a ten year period.

For some tally trees, e.g. ingrowth trees and ongrowth trees, past periodic measurements of DBH growth may not be available. For such trees, an equation is needed to estimate basal area growth from current tree characteristics. If other tally trees with measured periodic growth are present on the plot, a simple power function of diameter is used, developed from tree measurement data from that plot. The model is:

$ADD^2 \quad = \quad b(D2)^c$, where: b and c are coefficients.

If other tally trees with remeasured growth are not present, a generalized squared diameter growth model is used. For example, our current model to estimate the change in annual squared diameter for conifers growing on plots without growth information is:

$ADD^2 \quad = \quad f(D^{.7}*CR^{1.8}, D^2/A, S_{1..i}*D^2/A, CC4*D^2/A)$

where: CR = crown ratio

$S_{1..i}$ = dummy variables for individual species

$CC4$ = dummy variables for tree social position

A = breast high age.

Squared DBH growth is calculated and stored for each tally tree during inventory compilation. The information is, therefore, readily available at any time for updating purposes.

Tree Height Growth

In the dense, tall forests of the Douglas-fir subregion, tree height growth cannot be measured accurately over a 10-year period. For this reason, we use site index curves as the basis for estimating periodic height growth. Specifically:

$DH_c \quad = \quad [(H2_m/H2_e)(H2_e-H1_e)x]$

where: DH_c = change in heights between measurements

c = years between measurements

$H2_m$ = measured height at second measurement

$H1_e,H2_e$ = hypothetical heights calculated from site equation for a tree of same species and age as the tally tree at each measurement occasion

x = calibration adjustment factor.

A sample of trees are measured for height growth at each measurement occasion. Although these repeat measurements provide an unreliable estimate of individual tree height growth, they collectively provide an unbiased estimate of average height growth, by species, for the trees measured. Thus, height growth is calibrated to inventory measurements as follows:

$x = \geq(H2_m\text{-}H1_m/n)/\geq(DH_c/n)$,

where: x = adjustment factor

$H1_m$ = height measured at the first measurement

n = number of tally tree measured at both occasions.

Modeling mortality

We discount the trees-per-acre expansion factor for each tally tree by it's probability of mortality during the update period. The probability of dying is, for each tally tree, a function of species, crown ratio, and age or breast high diameter. We estimate these probabilities either from equations based on the model presented by Hamilton and Edwards (1976), or by developing mortality rates by species, crown ratio, and DBH class.

Updating for the effects of harvest activities

The Oregon Department of Forestry and Washington Department of Natural Resources compile annual statistics on log production in Oregon and Washington (for example: Oregon Department of Forestry 1988). These data are complete enumerations of aggregate harvested volumes by county and broad owner group (forest industry, national forest, etc.), but they provide no information on the species, size or specific location of the harvested trees.

We accept these harvest statistics as the best available. Therefore, our problem is not one of estimating the total aggregated harvest but one of allocating that harvest in a way that reflects it's impact on the forest conditions described by our inventories. An acceptable solution to this problem must allow us to compile updated inventories that provide the same statistical summaries required of our regular inventories. To produce these summaries means we must simulate the effects of harvest on each plot and each tally tree. The technique must be low in cost, a requirement that eliminates revisiting plots to monitor for harvest changes since last inventory.

The complexity of updating for the effects of harvest depends, in part, on the number and characteristics of the harvesting methods that are common within the region to be updated. In the Douglas-fir subregion, only one kind of harvest activity--clearcutting--is important in terms of volume removed, the acreage harvested or the stand conditions affected. In western Oregon, for example, we estimate from FIA inventories that 96-percent of the harvest on non-National Forest lands in the period 1973 to 1986 was removed by clearcutting (unpublished estimate by authors). That only clearcutting is important means we need not develop a multinomial estimator or individual estimators to estimate the occurrence for several types of harvest on FIA timberland plots. For the acres estimated as clearcut, we can simply assume that all trees were removed or killed; were partial cutting important, we would be faced with the additional and more complex problem of estimating which trees were harvested plot-by-plot.

Our approach to update for harvest effects in the Douglas-fir subregion is to develop two components: (1) an estimator to estimate, for specified update periods of one to 10 years, the annual probability of clearcut harvesting on FIA timberland plots with volume present; and (2) an algorithm to adjust FIA plot and tree files--given the plot-level probabilities from the first component--for the effects of known harvest levels reported for the specified period.

In estimating the occurrence of harvest, we are focusing on a probabilistic approach instead of one that is deterministic. While a deterministic approach provides a black-and-white decision--a plot is either cut or not cut during an update period--conceptually, the decision implies that all acres that the plot represents--the plot's area expansion factor--are also cut or not cut. In reality, not all of these acres are apt to be harvested during an update period (<10 years), but only some fraction, a proportion that is equivalent to the probability of harvest occurring in the period. Thus, our strategy is to develop regression equations which estimate the annual probability of harvest by plot in which the probability is a function of characteristics measured or observed during prior inventories.

While many probability models exist, two, the probit (normal) and the logit (natural logarithm) cumulative probability functions, have been widely applied in forestry and other fields to model event probabilities. Both are asymptotic at 0 and 1, a desirable characteristic when estimating probabilities. Both have curvilinear response surfaces that are not statistically significantly different (Capps and Kramer 1985). The logit model, however, is computationally simpler to fit.

The general form of the logistic harvest model is:

$$P = 1 / 1 + \exp(-[b_0 + b_1x_1 + b_2x_2 + \ldots + b_ix_i]),$$

where: P = the annual probability of harvest,

b_i = $_i$th nonlinear regression coefficient (model parameters), associated with independent variables

x_i = $_i$th independent variable.

Logit models have been applied by several researchers to model the annual average probabilities of tree mortality (Hamilton 1974; Hamilton and Edwards 1976; Monserud 1976; Buckman 1979; Buckman et. al. 1983). To develop an estimator of harvest probabilities, the model is fitted from an set of observations--FIA timberland plots with volume present--for which the dependent variable--the occurrence of harvest between two successive inventories--is dichotomous: 0 if clearcut harvest has not occurred and 1 if it has. The coefficients of the independent variables are calibrated from data collected at the first of these inventories.

A key problem in developing an estimator is that the period between inventories often varies among observations in the development set of data; for example, plots in western Oregon vary from 8 to 12 years for the period between latest and prior inventories. Two solutions are possible. The first is to annualize the observation set by dividing the dependent variable for each plot by the number of years between inventories; in effect this treats the problem as one of simple interest, a technique employed by Hamilton and Edwards (1976) in modeling tree mortality. The other is to treat the occurrence of harvest as a compound interest phenomenon in which the general model is modified to the form:

$$P = 1 / \{1 + \exp(-[b_0 + b_1x_1 + b_2x_2 + \ldots + b_ix_i])\}^n$$

in which "n" is the is the length of the remeasurement period. Monserud (1976) and Hamilton (in press) have applied this technique to model annual probabilities for tree mortality. Hamilton observed that this second solution is always more technically correct. In comparing the two methods, he found that changing from the first to the second solution during development of a mortality model did not change the variables included in his model but did eliminate most of the unexplained underestimation of mortality rates present using the first method. Hamilton noted that when annual mortality rates are low (0.5%) and measurement intervals range from 2 to 10 years, the two solutions provided similar models. In updating for harvest in western Oregon and Washington, treating the occurrence of harvest as a compound interest phenomenon may be the better choice since the average annual rate of harvest associated with older, high volume stands is relatively high, perhaps 2 to 3-percent a year.

The second component--an algorithm for updating plot and tree files--is based on partitioning the area expansion factors for each timberland plot estimated to have a probability of harvest greater than 0 in an update period. Prior to partitioning, the estimated annual probabilities of harvest for each plot are compounded, given the number of years in the update period, to periodic probabilities. Partitioning is accomplished by multiplying the expansion factor for each plot by it's periodic probability of harvest. For example, if a timberland plot represents 7,400 acres in the latest inventory and the periodic probability of harvest for the plot is .3, then 2,220 acres are estimated as clearcut and 5,180 acres, uncut.

In compiling updated statistics, the tree list for each timberland plot is grown forward and adjusted for mortality using the techniques described earlier. The updated tree list for each plot is expanded by the acreage estimated as uncut (5,180 acres in the example). For the acres estimated as harvested, the tree list is assumed empty. To estimate characteristics of harvest--volume by species, diameter class, etc--one would expand the appropriate trees on the tree list at last inventory by the number of cut acres.

To reconcile aggregate estimates of harvested volume with independently reported harvest statistics requires adjusting the harvested acres estimated for each timberland plot in the inventory by the following ratio R:

$$R = X/Y,$$

where: X = the aggregated harvested volume for the update period independently reported, and

Y = the aggregated harvested volume for the update period compiled from the inventory using harvest probabilities.

The adjusted harvested acreage for each plot is subtracted from the plot's original expansion factor (7,400 acres in the example) to obtain an adjusted estimate of uncut acres. These adjusted plot-level estimates for harvested and undisturbed acreages are used to compile updated inventory statistics.

CONCLUSION

We are developing techniques for updating FIA inventories in the Douglas-fir subregion that meet local, regional and national requirements for current information about forest resources. Our basic assumption is that we can estimate changes in forest conditions over short periods of time by extrapolating recent rates of change. These rates are detected through successive measurements and observations of the same permanent FIA plots on which the inventories are based. While our methods are well suited for updating inventories in the subregion, they may work less well in regions where the components of forest change are more variable.

LITERATURE CITED

Arney, James D. 1985. A modeling strategy for the growth projection of managed stands. Canadian Journal of Forest Research. 15:511-518.

Buckman,R.G. 1979: Mortality functions. In: A generalized forest growth projection system applied to the Lake States region. Gen. Tech. Rep. NC-49. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station: 47-55.

Buckman, R.G.; Pederson, S.P.; Walters, N.R. 1983. A tree survival model with applications to species of the Great Lakes region. Canadian Journal of Forest Research. 13:601-608.

Capps, O.; Kramer, R.A. 1985. Analysis of food stamp participation using qualitative choice models. American Journal of Agricultural Economics. 67:49-59.

Curtis, Robert O., Clendenen, Gary W., DeMars, Donald J. 1981. A new stand simulator for coast Douglas-fir. DFSIM user's guide. Gen. Tech. Rep. PNW-128. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station. 79 p.

Hamilton, David A. 1974. Event probabilities estimated by regression. Res. Pap. INT-152. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 20 p.

Hamilton, David A.; Edwards, B.M. 1976. Modeling the probability of individual tree mortality. Res. Pap. INT-185. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 22 p.

Hamilton, David A. [in press]. Spiffy little things: elements in the evolution of a mortality model. Canadian Journal of Forest Research.

Hester, Arlene S., Hann, David W., Larsen, David R. 1989. Organon: Southwest Oregon growth and yield model user manual. Version 2.0. Corvallis, OR: Oregon State University, Forest Research Laboratory. 59 p.

MacLean, Colin D. [in press]. Changes in area and ownership of timberland in western Oregon: 1961-1986. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

MacLean, Colin D. and Scott, Charles T., 1988. Interpolating current annual growth from two d.b.h. measurements. In: Proceedings, IUFRO conference, Forest Growth Modelling and Prediction; 1987 August 23-27; Minneapolis, MN. Gen. Tech Rep. NC-120. St. Paul, MN: U.S.Department of Agriculture, Forest Service, North Central Forest Experiment Station; 1118-1123.

Monserud, R.A. 1976. Simulation of forest tree mortality. Forest Science. 22:438-444.

Oregon Department of Forestry. 1988. Oregon timber harvest report 1987. Salem, OR. 2 p.

Shifley, Stephen R.; Fairweather, Stephen E. 1983. Individual tree diameter growth and mortality models for western Oregon. In: Proceedings, Renewable Resource Inventories for Monitoring Changes and Trends; an International Conference; 1983 August 15-19; Corvallis, OR. Corvallis, OR: Oregon State University; p.591-595.

Smith, Brad. 1989. Updating forest inventories: the North Central experience. These proceedings.

Kalman Filter to Update Forest Cover Estimates

Raymond L. Czaplewski
Mathematical Statistician
USDA Forest Service
Rocky Mountain Forest and Range Experiment Station
240 West Prospect Road
Fort Collins, Colorado 80526 U.S.A.

ABSTRACT

The Kalman filter is a statistical estimator that combines a time-series of independent estimates, using a prediction model that describes expected changes in the state of a system over time. An expensive inventory can be updated using model predictions that are adjusted with more recent, but less expensive and precise, monitoring data. The concepts of the Kalman filter are explained with a simple, hypothetical example of estimating percent forest cover over time, using remote sensing and field plots from a forest inventory.

INTRODUCTION

The Kalman filter is a composite estimator, which combines two independent estimates at a time, each of which is weighted inversely proportional to its variance. Gregoire and Walters (1988) note that composite estimators are widely used in forestry, including sampling with partial replacement (e.g., Ware and Cunia 1962). Green and Strawderman (1986) and Thomas and Rennie (1987) show how a composite estimator may be used to combine independent estimates of stem density, basal area, or wood volume.

In the Kalman filter, one of the independent estimates is a current estimate or monitoring measurement (e.g., remotely sensed data or midcycle update); the other is a previous estimate, or forest inventory, that is updated for expected changes over time using a deterministic prediction model. Variance for this updated estimate includes effects of (1) errors in the previous forest inventory that are propagated over time, and (2) model prediction errors between the previous and current estimates. Errors in a composite estimate are typically less than errors in either prior estimate alone.

EXAMPLE OF A COMPOSITE ESTIMATE

First, consider the problem of estimating percent forest cover in Fig. 1. Make an ocular estimate of the percent forest in Fig. 1, now. Then record your estimate in Table 1.

Nine other photointerpreters independently made ocular estimates for this same image, and their results are recorded in Table 1. Assume each ocular estimate has no bias or sampling error, and each has an identical distribution of measurement errors (i.e., Fig. 1 is imperfectly censused, with nine independent replicates). The mean of these ocular measurements ($x_1 = 56.0\%$ from Table 1) is the first estimate of percent forest cover in Fig. 1, with variance of the mean $\mathrm{Var}(x_1) = 3.11\%\%$ (variance units are $\%^2$, denoted as $\%\%$)

Next, consider a second estimate of percent forest cover in Fig. 1 using error-free classification of 400 temporary plots. There are 204 forested plots, producing the estimate $y_1 = 51.0\%$. Using the binomial distribution, the estimated sampling variance is $\mathrm{Var}(y_1) = (51.0\%)(49.0\%)/400 = 6.25\%\%$, which produces the approximate 95% confidence interval $\mathrm{CI}_{95\%}(y_1) = 51.0 \pm 4.9\%$.

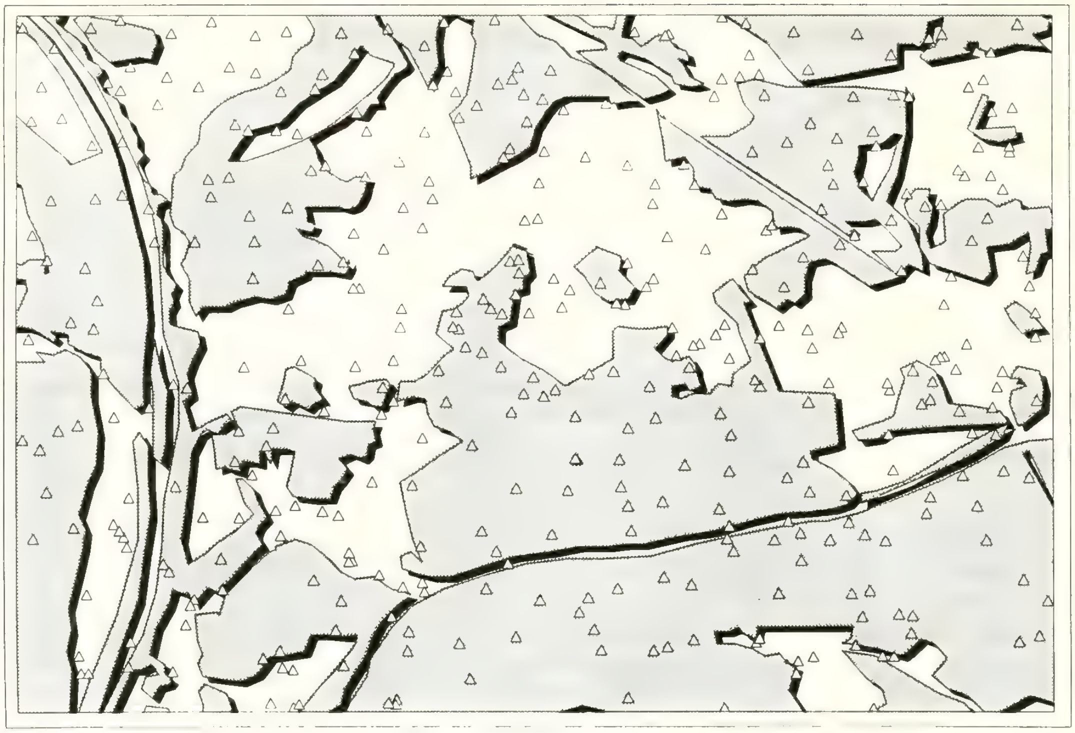


Figure 1. Hypothetical example requiring an estimate of the percent forest cover (shaded area in this image) at time $t=1$. You are requested to make an ocular estimate of percent forest, and record your estimate in Table 1. Subsequently, 400 randomly located plots are independently classified into forest or nonforest.

Table 1. Independent ocular estimates (photointerpretations) of percent forest in Fig. 1 by nine (or more) different, but unbiased and equally skilled, observers. These are pooled into a group estimate. Recompute the mean, variance among interpretations, and variance of the mean, using your ocular estimate of percent forest cover in Fig. 1 ($n=10$).

Observer	Ocular Estimate (x)	x^2	Estimator Calculations ($n=9$)
1	60	3600	Group mean $x_1 = 504\%/9$
2	55	3025	$x_1 = 56\%$
3	50	2500	
4	55	3025	Variance among photointerpretations
5	52	2704	$=[28448\%\%-(504\%)^2/9]/8$
6	50	2500	$=28\%\%$
7	65	4225	
8	62	3844	Variance of mean$=28\%\%/9$
$n=9$	55	3025	$\mathrm{Var}(x^1)=3.11\%\%$
Subtotal	504%	28448%%	
			Approximate 95% confidence interval
Yours	___	_____	$x_1 \pm 1.96(3.11)^{1/2}=56.0\pm3.5\%$
Total($n=10$)			

The estimate $x_1=56.0\%$ from Table 1 can be combined with the sample estimate from Fig. 1 ($y_1=51.0\%$) to produce a new composite estimate x^*_1, as shown in Fig 2. x_1 is weighted more heavily than y_1 because $\mathrm{Var}(x_1)=3.11\%\%$ for replicated ocular measurement error is less than $\mathrm{Var}(y_1)=6.25\%\%$ for sampling error from 400 point plots:

$$\begin{aligned} x^*_1 &= [A_1\ y_1] + [(1 - A_1)\ x_1] \\ &= [(0.33)\ 51.0\%] + [(0.67)\ 56.0\%] \\ &= 54.4\% \end{aligned}$$

$$\begin{aligned} A_1 &= \mathrm{Var}(x_1)/[\mathrm{Var}(x_1)+\mathrm{Var}(y_1)] \\ &= 3.11\%\% \ / \ (3.11\%\% + 6.25\%\%) \\ &= 0.33 \end{aligned}$$

Weight (or shrinking coefficient) A_1 can be derived using maximum likelihood, minimum variance, or empirical Bayes theory. The expected variance of the composite estimate $\mathrm{Var}(x^*_1)$ is based on an elementary theorem in mathematical statistics for the variance of a linear transformation Y of independent random variables X_a and X_b:

$$\mathrm{Y} = (\mathrm{a}\ X_a) + (\mathrm{b}\ X_b)$$

$$\mathrm{Var(Y)} = \mathrm{a}^2\ \mathrm{Var}(X_a) + b^2\ \mathrm{Var}(X_b)$$

Applying this theorem to the composite estimator x^*_1:

$$\begin{aligned} \mathrm{Var}(x^*_1) &= [A_1{}^2\ \mathrm{Var}(y_1)] + [(1 - A_1)^2\ \mathrm{Var}(x_1)] \\ &= [(0.1089)\ 6.25\%\%] + [(0.4489)\ 3.11\%\%] \\ &= 2.08\%\% \end{aligned}$$

This estimator of $\mathrm{Var}(x^*_1)$ is also given by Gregoire and Walters (1988). The variance of the composite estimate is smaller than either of the two independent estimates (Green and Strawderman 1986), as illustrated in Fig. 2; the approximate 95% bounds on estimation error for the composite estimate x^*_1 are ±2.8%, compared to ±3.5% for the mean ocular estimates, and ±4.9% for the sample estimate using 400 plots.

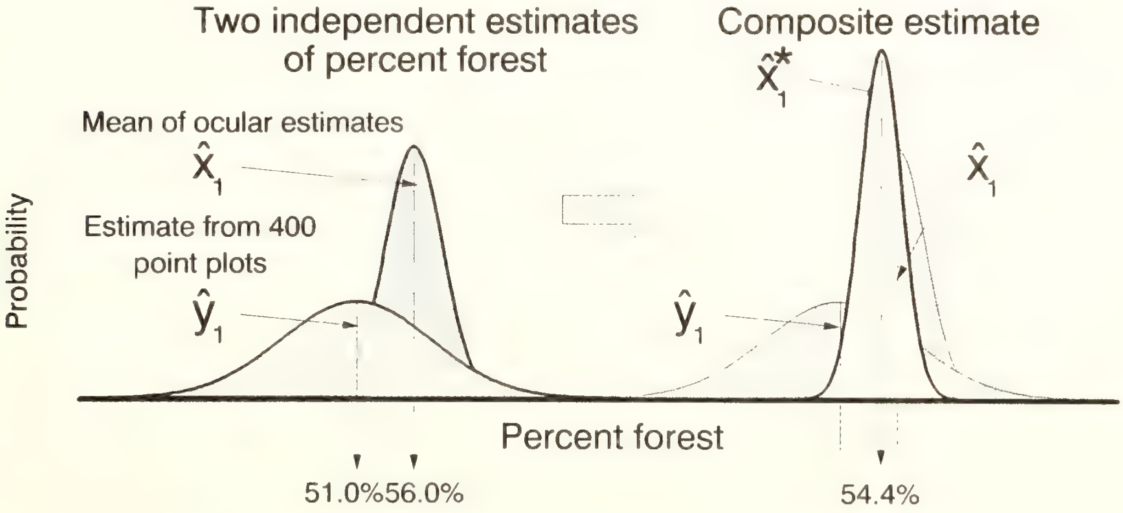


Figure 2. Probability densities for estimates of percent forest cover from mean ocular estimates ($x_1=56.0\%$) and 400 point plots ($y_1=51.0\%$). These are weighted inversely proportional to their variances, and combined into the composite estimate ($x^*_1=54.4\%$).

CHANGES OVER TIME

Periodic forest inventories, inventory updates, or forest monitoring programs seek estimates of forest condition after changes have occurred. Fig. 3 shows the condition of forest cover at time $t=2$, after clearcuts have changed the hypothetical example in Fig. 1 at time $t=1$. Make another ocular estimate of percent forest cover in Fig. 3, now. Then, record your answer next to Fig. 3.

In Fig. 3, 200 temporary plots are independently classified to estimate percent forest cover at $t=2$; 90 plots have forest cover, for an estimate $y_2=45.0\%$, with $\mathrm{Var}(y_2)=(45.0)(55.0)/200=12.38\%\%$, and $\mathrm{CI}_{95\%}(y_2)=45.0\pm6.9\%$.

Another estimate of percent forest x_2 at time $t=2$ can be made from prior estimate $x^*_1=54.4\%$, given an estimated rate of change. It is predicted, from historical trends or a deterministic model, that 5% of all forest cover in Fig. 1 is clearcut between $t=1$ and $t=2$; therefore, the expected percent forest at $t=2$ is $x_2=0.95(x^*_1)=0.95(54.4\%)=51.7\%$.

If this linear transformation x_2 of x^*_1 is a perfect prediction model, variance of the propagated estimation error at $t=2$ is simply $\mathrm{Var}(x_2)=(0.95)^2\mathrm{var}(x^*_1)=(0.90)2.08\%\%=1.88\%\%$. However, models are imperfect; prediction errors, denoted w, occur; and linear transformation $x_2=0.95x^*_1+w$ is a more realistic prediction model. If prediction error is unbiased, i.e., the expected value $\mathrm{E}[w]$ is zero, then the updated estimate x_2 is unaffected.

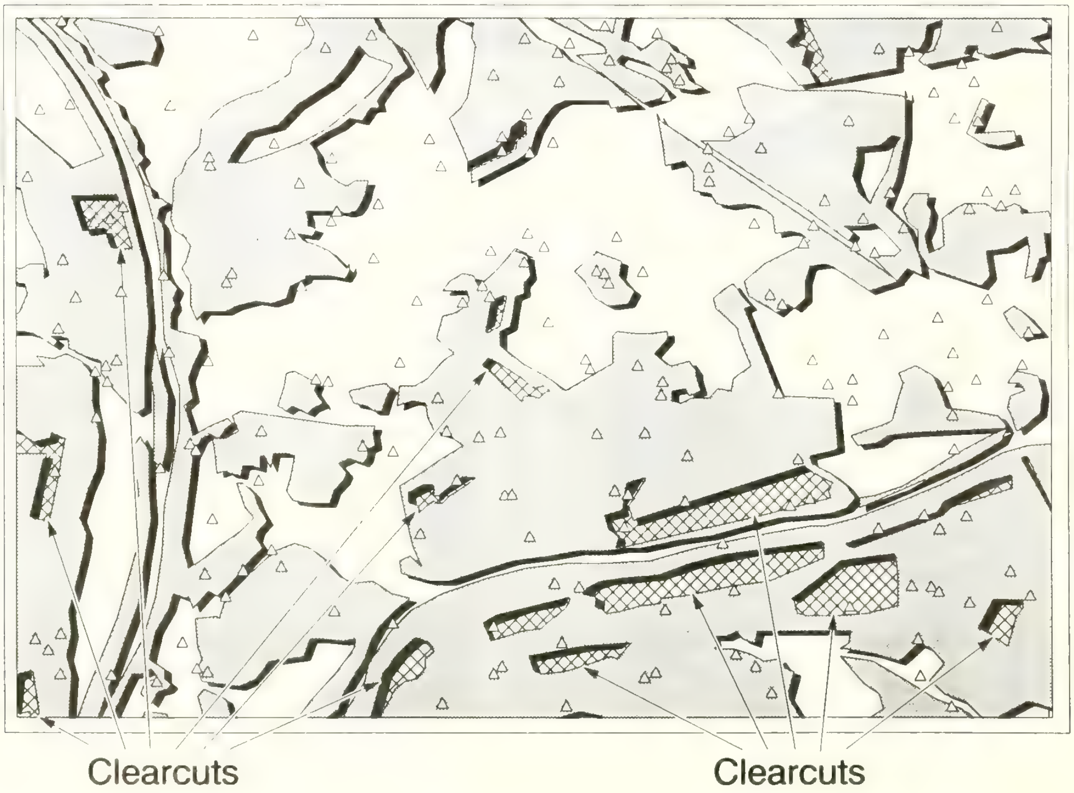


Figure 3. Between times $t=1$ and $t=2$, clearcuts reduce forest cover. In a monitoring system, 200 point plots are used at $t=2$ to independently estimate percent forest cover.

If prediction errors w are independent of errors for x^*_1, then $\mathrm{Var}(x_2)=(0.95)^2\mathrm{Var}(x^*_1)+\mathrm{Var}(w)$. If $\mathrm{Var}(w)$ is assumed to be 1.00%%, then $\mathrm{Var}(x_2)=1.88\%\% + 1.00\%\% = 2.88\%\%$, which yields an updated estimate x_2 with $\mathrm{CI}_{95\%}(x_2)=51.7\pm 3.3\%$.

The Kalman filter is merely a different linear transformation (x^*_2) of the estimates y_2 and x_2, with weights (A_2) inversely proportional to their variances (Fig. 4):

$$\begin{aligned} A_2 &= \mathrm{Var}(x_2)/[\mathrm{Var}(x_2)+\mathrm{Var}(y_2)] \\ &= 2.88\%\% \;/\; (2.88\%\% \;+\; 12.38\%\%) \\ &= 0.19 \end{aligned}$$

$$\begin{aligned} x^*_2 &= [A_2\, y_2] \;+\; [(1 - A_2)\, x_2] \\ &= [(0.19)\; 45.0\%] \;+\; [(0.81)\; 51.7\%] \\ &= 50.4\% \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(x^*_2) &= [A_2^{\,2}\; \mathrm{Var}(y_2)] \;+\; [(1 - A_2)^2\; \mathrm{Var}(x_2)] \\ &= [(0.0361)\; 12.38\%\%] \;+\; [(0.6561)\; 2.88\%\%] \\ &= 2.34\%\% \end{aligned}$$

Approximate 95% bounds on estimation errors for the Kalman composite estimate are ±3.0%, compared to ±6.9% for the sample estimate y_2, and ±3.3% for model-updated estimate x_2.

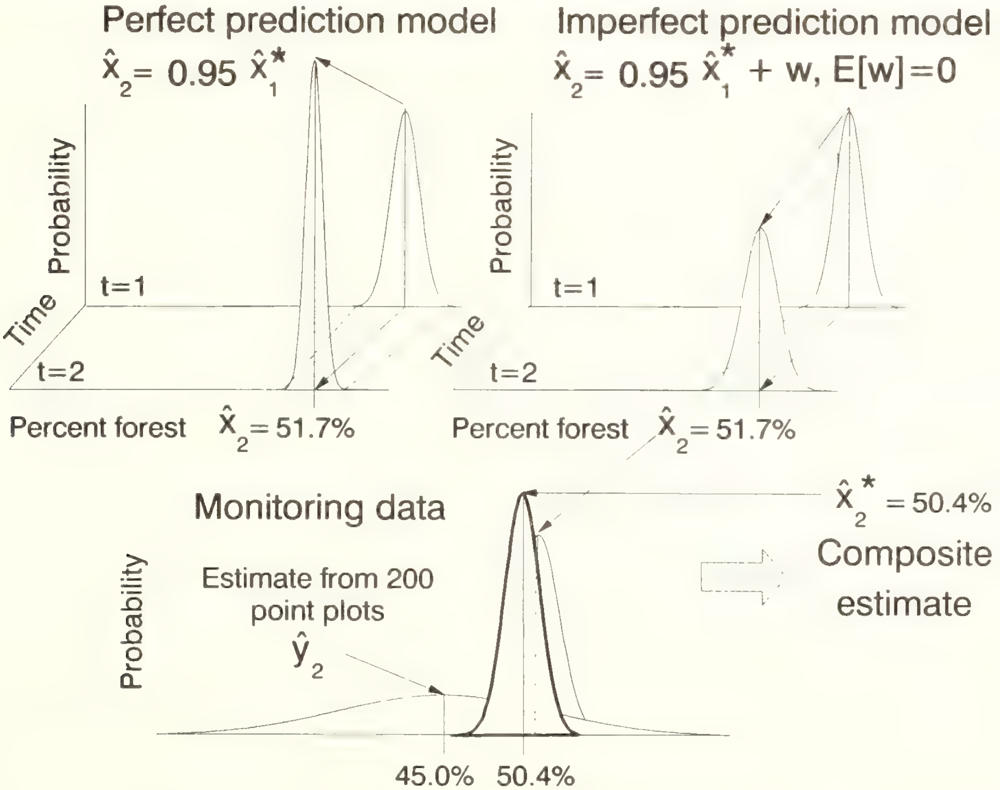


Figure 4. Probability densities for estimates of percent forest cover at time $t=2$, made from a prior estimate at time $t=1$, given model $x_2=(0.95)x^*_1=51.7\%$ (5% of the forest is expected to be clearcut between $t=1$ and $t=2$). Given a perfect prediction model, only the estimation error at $t=1$ is propagated to $t=2$. More realistically, the model is imperfect, and an unbiased prediction error (w) also occurs. An independent estimate y_2 is available from the 200 plots in Fig. 3. The Kalman filter combines these two independent estimates into a composite estimate (x^*_2), with weights inversely proportional to their variances.

TIME-SERIES OF MONITORING DATA

The Kalman filter is usually applied to a time series of monitoring measurements. With each new monitoring measurement, a composite estimate is made, which serves as new initial conditions for the next deterministic prediction (e.g., Fig. 5, year 4). Monitoring measurements can adjust predictions from a simple linear model for trends that are truly nonlinear, but are not well quantified. Precise data from the past can improve current estimates using less precise, but more recent, monitoring data. The Kalman filter can combine monitoring data from many sources, e.g., remote sensing or severance tax records.

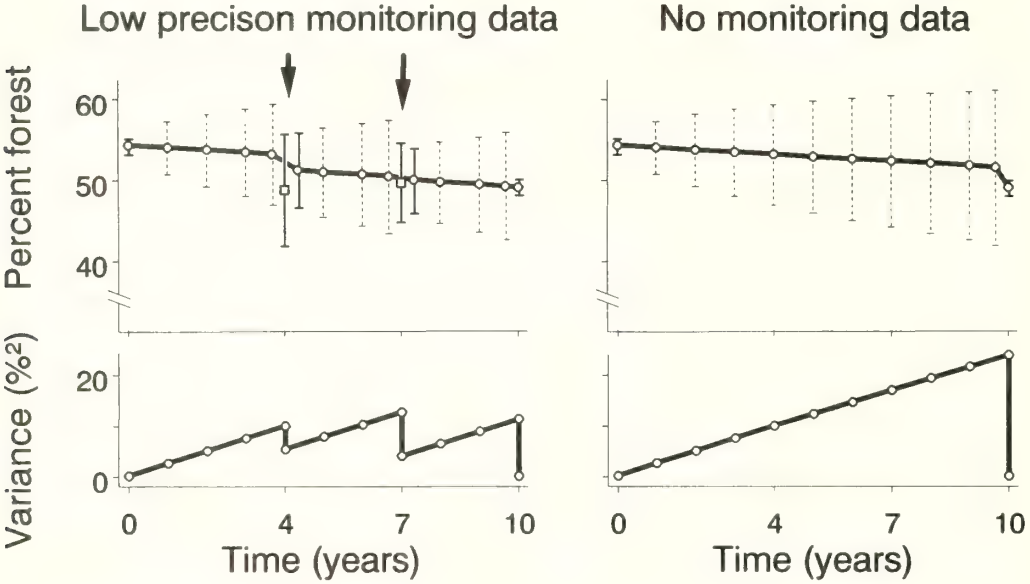


Figure 5. Kalman estimates and approximate 95% confidence intervals. Forest inventories were conducted in years 0 and 10; monitoring data were gathered in years 4 and 7. A time series of relatively imprecise (i.e., inexpensive) monitoring data can prolong utility of a previous, more expensive forest inventory.

TUTORIAL

To better understand the basic functions of the Kalman estimator, recompute the estimates above, using your ocular estimate of percent forest in Fig. 1, which you recorded in Table 1, as follows:

1. Compute a new mean of the ocular estimates x_1, and its variance of the mean $\mathrm{Var}(x_1)$, using $n=10$ in Table 1.

2. Recompute x^*_1, which combines this new x_1 with the sample estimate from the 400 point plots: $y_1=51.0\%$, $\mathrm{Var}(y_1)=6.25\%\%$.

3. Apply the prediction model $x_2=(0.95)x^*_1$, and the estimator for $\mathrm{Var}(x_2)$, where $\mathrm{Var}(w)$ is given as $1.00\%\%$.

4. Combine this updated estimate (x_2) with the sample estimate at $t=2$ [$y_2=45.0\%$, $\mathrm{Var}(y_2)=12.38\%\%$] to compute x^*_2 .

The resulting composite estimate (x^*_2) is also the Kalman filter estimate. Compute an estimate of percent forest at $t=3$, given that 470 plots were forested in a sample of 1000 point plots at $t=3$, and using the same model for change (i.e., 5% of forest cover is clearcut between $t=2$ and $t=3$, with no regeneration). Repeat 1 to 4 above using only your ocular estimate (x_1) of Fig. 1, with $\mathrm{Var}(x_1)=28\%\%$ from Table 1, and vary the estimate of model prediction error dispersion, e.g., $\mathrm{Var}(w)=5\%\%$ or $10\%\%$. Assume ocular estimates are biased, and the true percent forest (x_1) is related to your ocular estimate (x_0) by the known linear model $x_1=0.85(x_0)+2\%$; recompute the composite estimate x^*_1 using only your ocular estimate (x_0) for Fig. 1. Assume the model that predicts x_1 from your x_0 is not precisely known, but is estimated by linear regression using a finite sample of reference sites ($n=100$) from a calibration experiment, in which residual mean square error is $30\%\%$, mean of the 100 ocular estimates x_0 is 50.5%, and sum of $(x_0\text{-}50.5\%)^2$ in the experiment is $84000\%\%$; recompute composite estimate x^*_1 using your calibrated ocular estimate x_1 of x_0. The true percent forest (usually unknown) is 54.68% at $t=1$ (Fig. 1), not including shadows, 49.65% at $t=2$ (Fig. 3), and 45.08% at $t=3$.

VERIFICATION

It is possible that two independent estimates disagree, or "diverge", in that neither estimate is likely given the other (Fig. 6). Contradictory estimates can be combined, but the result can be biased. Contradictions are probably caused by problems in estimating the error distribution of (1) the current monitoring measurement, or (2) the past estimate that is updated by a deterministic prediction model.

Concerning (2) above, it is very difficult to estimate the variance of model prediction errors $\mathrm{Var}(w)$; the estimate $\mathrm{Var}(w)=1.00\%\%$ in the hypothetical example above is quite arbitrary. Accurate estimates of $\mathrm{Var}(w)$ would require known differences between model predictions and the true, but unknown, state of the system. As an alternative, adaptive filters modify initial, but inaccurate, estimates of $\mathrm{Var}(w)$ until disagreements are within acceptable bounds (Fig. 6), often using a time series of residuals $y_t\text{-}x_t$ (Sorenson 1985). Some Bayesian techniques assume model prediction errors are biased, and choose a weight or shrinking coefficient that minimizes a risk function.

It is also possible that bias exists in the measurement model. For example, different photointerpreters are rarely unbiased, and never have identical error distributions. Calibration experiments for each interpreter are needed to more realistically model such measurement errors.

The composite estimator assumes the two prior estimates are independent; however, their errors might be correlated. If your estimate of percent forest in Fig. 3 was influenced by any estimate of Fig. 1, then your measurement errors are correlated. The Kalman filter can treat correlated errors, but covariance estimates are required. Temporal or other patterns in standardized residuals from the Kalman filter can be used to detect such problems.

Negative variance estimates and asymmetrical covariance matrices can occur with the Kalman filter, especially when multivariate state vectors x_t are estimated and there are large differences between covariance matrices for multivariate x_t and y_t. Solutions to these numerical problems abound in the engineering literature, with the "square root" filter being frequently employed (Maybeck 1979).

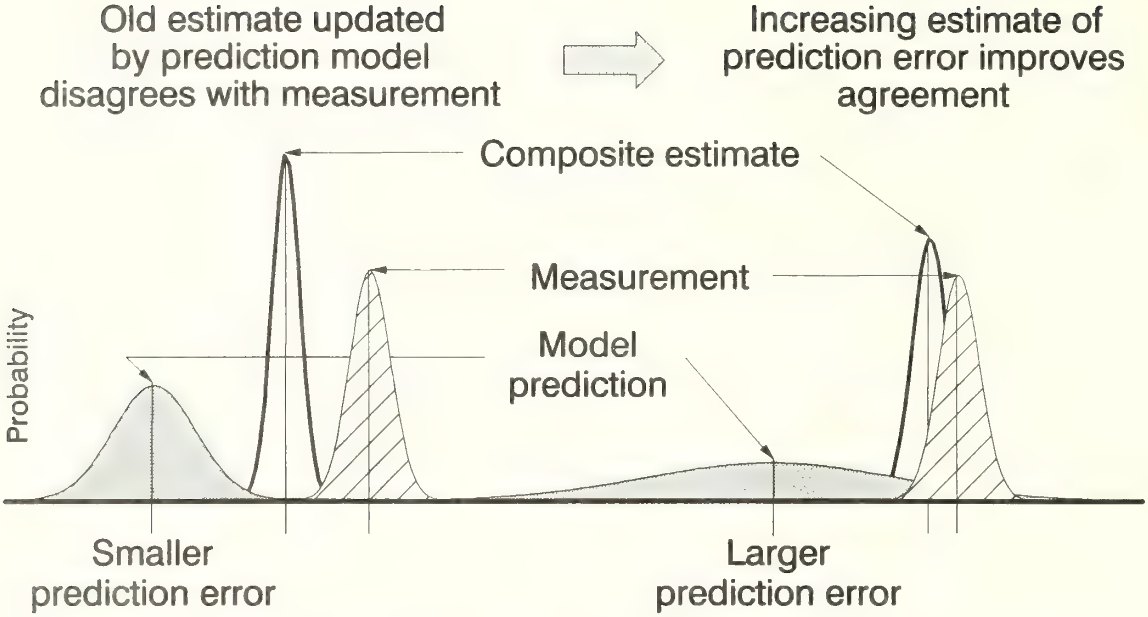


Figure 6. Predicted probability densities for two independent estimates (measurement y_t and model prediction x_t) that disagree; the residual difference between the two estimates is unlikely. This is probably caused by an unanticipated bias in the estimated error distribution for the measurement or model update. Adaptive filters assume that the estimated variance of model prediction errors Var(w) is inaccurate, and Var(w) is changed until the disagreement is within acceptable bounds.

CONCLUSIONS

The Kalman filter can be used to combine a time series of forest inventories, updates, or monitoring estimates using a model of expected change in forest condition over time. Many independent sources of data and knowledge can be combined, as discussed by Czaplewski et al. (1988). This can afford substantial efficiencies, and more timely estimates, compared to any one source of information by itself.

The Kalman filter is not complicated, but its application can be challenging. Accurate models for measurement and prediction errors are needed. Model formulation, parameter estimates, and assumptions must be verified; if residuals fail verification tests, then sound technical judgment is required to diagnose and cure any problems. However, solution of such problems can improve analysts' understanding of forest dynamics and the measurement processes.

REFERENCES

Czaplewski, R. L., Alig, R. J., and Cost, N. D. 1988: Monitoring land/forest cover using the Kalman filter: a proposal. In proceedings Forest Growth Modelling and Prediction, Minneapolis, MN, U.S.A. pp. 1089-1096. USDA Forest Service GTR NC-120, Saint Paul, MN. 1153p.

Green, E.J. and Strawderman, W.E. 1986: Reducing sample size through the use of a composite estimator: an application to timber volume estimation. Can. J. For. Res., vol. 16, no 5, pp. 1116-1118.

Gregoire, T.G. and Walters, D.K. 1988: Composite vector estimators derived by weighting inversely proportional to variance. Can. J. For. Res., vol. 18, no. 2, pp. 282-284.

Maybeck, P.S. 1979. Stochastic Models, Estimation and Control, Vol. 1. Academic Press, New York. 423pp.

Sorenson, H.W. 1895. Kalman filtering: Theory and Applications. IEEE Press Inc., New York. 457pp.

Thomas, C.E. and Rennie, J.C. 1987: Combining inventory data for improved estimates of forest resources. South. J. Appl. For., vol. 11, no. 3, pp. 168-171.

Ware, K.D. and Cunia, T. 1962: Continuous forest inventory with partial replacement. For. Sci. Monogr. No. 3.

Integrating Multiple Value Forest Surveys Into Timber Survey

Moderator: Victor A. Rudis

Richard E. Rice, The Importance of Forest Inventories to the Conservation of Biological Diversity

Frances A. Hunt, The Forest Inventory: Incorporating Fisheries and Wildlife Values

Curtis H. Flather, Linda A. Joyce, and Rudy M. King, Linking Multiple Resources Analyses to Land Use and Timber Management: Application and Error Considerations

Linda A. Joyce and Curtis H. Flather, Using Land Base Inventories to Access Range and Wildlife Resources Nationally

Victor A. Rudis, Multiple Value Forest Surveys in the Midsouth States

Noel D. Cost, Expanding Conventional Timber Inventories into a Broader, Multiresource Inventory in the Southeastern States

Renee O'Brien, Developing Integrated Forest Surveys for the Rocky Mountain States

J. David Born, Nonindustrial Wood Resources in the Rocky Mountain States

Bert Mead and Vernon J. LaBau, Multiresource Inventories in Alaska

Janet L. Ohmann, Multiresource Inventories in the Pacific Coast States—Progress and Future Direction

Robert T. Brooks, State-of-the-Art Methodology of Forest Inventory: Wildlife Habitat Assessment in the Northeastern United States

Workshop 2

The Importance of Forest Inventories to the Conservation of Biological Diversity

Richard E. Rice
Resource Economist
Resource Planning and Economics Department
The Wilderness Society
1400 Eye Street, N.W., Washington, D.C. 20005

ABSTRACT

Conserving biological diversity is becoming an increasingly important focus of national and international concern. The U.S. Forest Service is uniquely situated to contribute to this effort. With 191 million acres in 42 states, the U.S. Virgin Islands, and Puerto Rico, most major ecosystems found in North America are represented somewhere on a national forest or national grassland. In addition, while all federal agencies are bound to protect threatened and endangered species, the Forest Service is alone in having a legal mandate to preserve biological diversity. The agency's inventory and analysis staff could play a particularly important role in advancing this mandate. Adequate inventories are necessary to assess the impact of forest management on biodiversity and to devise and implement effective conservation strategies. To date, however, resource inventories have focused primarily on information of interest to commercial timber production. Current threats to diversity, the implications of failing to counter those threats, and the potential the Forest Service has to contribute to the conservation of diversity, all argue for a prompt and vigorous expansion of the agency's focus. A proposal is made to establish an independent conservation inventory that would emphasize priority species, habitats, and threats to biological diversity.

INTRODUCTION

I was invited here today to discuss The Wilderness Society's view of priorities for nontimber forest inventories. As many of you know, The Wilderness Society has a long-standing interest in the management and protection of our nation's forests. The list of nontimber resources of interest to the Society is a long one and I have no intention of trying to cover the waterfront in my remarks today. Instead, I would like to focus on a subset of those resources — known collectively as biological diversity — that I feel deserve priority attention from the Forest Service in general and from inventory personnel in particular.

Biological diversity is a term that has received a great deal of attention of late. It has been the source of more than its share of controversy, and as often as not, misunderstanding. To avoid as much misunderstanding as possible this morning, I would like to take a moment to outline what biological diversity is and why I feel that it is essential that the Forest Service play an active role in promoting its conservation in the United States.

BIODIVERSITY AND THE ROLE OF THE U.S. FOREST SERVICE

In its broadest sense, biological diversity is the diversity of life. Ecologists who study the patterns and processes of life tend to focus on diversity at three separate levels: genetics, species, and ecosystems (Norse *et al.* 1986, Crow 1988, Wilcove 1988). Each level interacts with the other in the form of a biological hierarchy — genes within species, species within ecosystems, and ecosystems within the broader landscape, or biosphere. All three levels *must* be considered in any comprehensive plan to conserve biological diversity.

Species diversity, the most familiar level, refers to the variety of plants, animals, fungi, and bacteria that inhabit the globe. Less obvious, perhaps, but no less important is the ***genetic diversity*** that exists among individuals and between populations within a given species. This level of diversity allows species and populations to adapt to changes in their environment and is the level at which the fundamental mechanisms of evolution are most apparent. In small or isolated populations, a lack of genetic diversity can lead to inbreeding, lowered "fitness," and extinction (Allendorf and Leary 1986; Ledig 1986; Ralls *et al.* 1986). Finally, ***ecosystem diversity*** is the diversity of distinctive groups or communities of species that inhabit different physical settings.

There are a number of reasons why I feel conserving diversity should be an important priority on the national forests. To begin, perhaps no land system in the United States offers more potential for the conservation of diversity than the National Forest System. The National Forest System is home to a third of all federally listed threatened and endangered species (USDA–Forest Service 1988b); it contains 191 million acres in 42 states, the U.S. Virgin Islands, and Puerto Rico; and most major ecosystems found in North America are represented somewhere on a national forest or national grassland (USDA–Forest Service 1988a).

In addition, the Forest Service is unique among federal land management agencies in having a legal mandate to preserve biological diversity. All federal agencies are bound to protect threatened and endangered species but only the Forest Service, through the National Forest Management Act of 1976, is required to "preserve and enhance the diversity of plant and animal communities" (36 C.F.R. 219.27(g)).

Finally, biodiversity on the national forests is being threatened and diminished by land-use practices and other factors on an alarming and unprecedented scale. While some threats to diversity, such as climate change and atmospheric deposition, are beyond the control of local managers, logging and other development activities are not. Managers should therefore have a clear idea of what diversity is and how their decisions are likely to affect it.

It is particularly important in this regard that managers maintain an awareness of the importance scale. Many forest management plans suggest that logging benefits diversity by creating a diverse age-class structure and a well distributed network of small forest openings (*eg.*, USDA–Forest Service 1986). Such a strategy may in fact lead to an increase in the number of species inhabiting a given area. Patch cuts create an abundance of forest edge which may be ideal for important game species such as deer, rabbit, and quail. But increased habitat for game species often comes at the expense of species most in need of protection. As a result, strategies aimed at maximizing *local* diversity often *lower* regional diversity.

A slightly different perspective may serve to reinforce this point (see Wilcove 1986). Estimates of the number of species on earth range from 1.5 to 30 million (Erwin 1982). We may never know the number precisely, but we can rest assured that whatever the number is, it will not increase significantly over the course of our lives, or the lives of our children, or even the lives of our children's children. The process of speciation operates on a time-scale of thousands or even millions of years. For all intents and purposes, then, the richness and variety of life on earth is fixed and we can no more increase it than we can increase the number of symphonies by Mozart or the number of statues by Michelangelo.

We are, however, extraordinarily adept at altering habitats so as to increase the populations of certain species and decrease the populations of others. In the process, we may either lower diversity or, if we are fortunate, not affect it significantly. But we will not increase it. The closest we can come to increasing diversity is to restore populations of species that have been extirpated in the wild. But even this will not *really* add to the diversity of life on earth. Moreover, were it not for the actions of man, the vast majority of these species would not have been threatened with extinction in the first place (USDA–Forest Service 1988a).

It is for this reason that not all species are equally important when it comes to conservation. Managers need to be particularly concerned with species that are prone to extinction and those that suffer most from current management practices. Habitat generalists, on the other hand, and species that tolerate a variety of management regimes, are likely to flourish without the assistance of the federal government. As a general rule, then, it is best for managers to manage locally for *regional* diversity (Crow 1988).

THE IMPORTANCE OF INVENTORIES TO CONSERVATION

How does all of this relate to forest inventories? Traditionally, forest inventories have focused on gathering information of interest to commercial timber production, and more recently, hunting, fishing, and other forms of recreation. While this is certainly worthy and important work, I would submit that the current threats to biological diversity, the impacts of failing to counter those threats, and the potential the Forest Service has to promote the conservation of diversity, all argue for a prompt and vigorous expansion of your focus.

The role of forest inventories in conserving diversity is a particularly important one. Among other things, inventories should serve as a baseline and foundation for other work — if you do not know what is out there, how it is distributed, how and where it is threatened, it will be difficult to implement or even devise effective conservation measures. Without adequate inventories, you cannot measure change over time; you will not know what gaps in knowledge exist, let alone be able to prioritize filling them; and you will have little idea of what and how management activities should be modified to ensure that the agency's conservation mandate is met. Simply put, good inventory data are the key to successful research, conservation, and management.

In recent years, Forest Inventory and Analysis units have, in fact, made some encouraging efforts to expand the focus of their work. The total vegetation inventories in the Pacific Northwest (Ohman 1989), work on red-cockaded woodpecker habitat on private lands in the Southeast (Flather 1986), and other wildlife-habitat modeling efforts are a few examples. I am concerned, however, that progress in the future will be hindered by a number of institutional obstacles.

First, your efforts seem to be limited from the outset to only those methods and procedures that can easily be incorporated into the timber survey (*eg.*, Brooks 1989). This is not always a constraint, of course, and in fact in many cases presents an opportunity to capitalize or leverage off an existing program. Nevertheless, the timber survey is clearly *not* an appropriate vehicle for all questions of interest and importance to conservation. At the very least, the species-habitat relationships used to derive non-timber information from timber surveys must be developed and verified first through direct species inventories.

Second, your current institutional structure seems to invite a lack of consistency in the type and quality of information collected in your surveys. At present, Forest Inventory and Analysis is primarily limited to surveys conducted on non-Forest Service lands. Comparable information from national forest inventories is often either nonexistent or not uniform from one forest to another. Moreover, your reliance on national forest *timber* surveys for non-timber information leaves you will little or no coverage in designated wilderness, national parks, or other areas of critical importance to conservation.

Finally, there is currently no formal institutional link between the collection of inventory information and management on the national forests. The absence of such a link has at least two important drawbacks. First, it lowers the chances that managers will actually use the information you develop, and second, it offers little opportunity for feedback from managers on how to develop more useful information.

RECOMMENDATIONS FOR BIODIVERSITY INVENTORIES

One means of addressing these shortcomings that I find particularly appealing would be to establish a separate authority or program for conservation inventories. This program could still leverage off of the timber survey where it was appropriate and effective to do so but would not be limited to the timber survey where better methods are available. Establishing a separate inventory would add to the program's visibility and help to promote support from Congress and the public. In addition, it would remove the damaging impression that timber is currently the real driving force behind your efforts.

To be effective, a separate conservation inventory would need to establish its own goals and objectives and ensure that adequate and uniform information was gathered from all forest ownerships. It would also require a focused agenda that was realistic and manageable from a technical and financial standpoint and yet responsive to the needs of conservation.

In this regard, there are a number of obvious priorities that I feel deserve broad attention. Among them, I would include at least the following three areas:

- *A targeted inventory of priority species.* It is far more important to know as much as possible about certain *targeted* species than to attempt to survey every species.

Included here would be rare, threatened and endangered species; top predators such as the gray wolf and grizzly bear; area-sensitive species, such as migratory songbirds that decline when forests are fragmented; plants and animals with limited ranges and/or low dispersal abilities; and species that are known to be adversely affected by current land-use practices.

Fortunately, there are already a number of excellent sources of information on the distribution and abundance of priority species. Two that come to mind are the various state heritage programs developed by The Nature Conservancy and the Forest Service's sensitive species program.

- *A targeted inventory of priority habitats, communities, and ecosystems.* Monitoring the distribution and condition of selected habitats or communities is an important complement to species-level inventories. Species, after all, depend on habitats, and for many species the disruption or elimination of habitat is what most threatens their survival.

Habitats or communities that are particularly species-rich, prone to disruption or subject to widespread degradation should be high priorities. Old-growth forests, for example, are important as much for the species that depend on them as for the pace at which they are being converted or degraded. Other examples include wetlands and riparian areas (particularly those grazed by cattle), and relic or isolated plant communities such as shale barrens and prairies. Centers of endemism such as the southern Appalachian Mountains, while not a habitat type *per se*, should also be given special attention.

- ***A targeted inventory of priority threats to biological diversity.*** Areas of particular importance in this regard are habitat fragmentation, the spread of exotic plants and animals, and climate change.

Habitat fragmentation is perhaps the single greatest threat to species diversity on the national forests (Norse *et al.* 1986). It is also one of the primary impacts of timber management (Harris 1984). Logging, road building, and the intentional creation of openings for game have transformed many formerly continuous expanses of natural forest into relatively small and increasingly isolated patches (Solheim *et al.* 1987). This, in turn, has given rise to a host of problems that, for many species, will eventually lead to extinction (*eg.*, Janzen 1986, Wilcove 1987). Ironically, the checkerboard pattern of harvests commonly used on the national forests to minimize the *visual* impacts of logging has probably *maximized* the amount of habitat fragmentation.

Assessing the degree of habitat fragmentation over broad geographical areas is probably most easily accomplished through a survey of roaded and roadless lands (*eg.*, Rudis 1988).

In many areas, competition from exotic plants and animals poses an additional threat to the viability of native flora and fauna. Examples include the wild boar in the southern Appalachians, and the gypsy moth and honeysuckle in eastern deciduous forests. Tracking the spread of these organisms may help managers identify methods of slowing their spread in the future (*eg.,* Craver 1982).

Finally, among the most serious long-term threats to species diversity is the prospect of a rapidly changing climate. Some scientists estimate that over the next half century, average annual temperatures in the United States may rise 2.7° to 8.1° F (1.5° to 4.5° C) (Mintzer 1987). Current projections of how forests will respond to this change are sobering indeed. Included are a reduction in the southernmost distribution of many northern hardwoods of up to 600 miles in the eastern U.S, an expansion of the northern range of southern pines into eastern Pennsylvania and New Jersey, and a reduction of mature, natural forest in the Great Lakes region and New England of from one-quarter to one-half their present volumes per unit of land (Winjum and Neilson 1989, Davis and Zabinski, In press). The precise timing and magnitude of many of these changes are unfortunately largely matters of speculation at present. It is clear, however, that if answers are to be had to many of these questions, long-term inventories will play an important role in helping to find them. And since larger networks and longer times-series generally make for stronger predictions, the sooner inventory sites are established the better.

Ideally, such a network would provide managers with one of the most important elements of a successful conservation strategy — time; time to plan for changes in advance, and an impetus and opportunity to develop a proactive rather than reactive approach to management.

More broadly, a conservation inventory should serve as a focal point for coordination among other research, conservation, and management activities. A primary objective of an inventory of roadless areas, for example, should be to identify opportunities for coordinated protection of large blocks of relatively intact forest — within forests, among groups of forests, and across jurisdictional boundaries with adjacent public and private lands.

Cross-jurisdictional inventories of public and private lands are also particularly well suited to making explicit the *regional* context within which conservation on the national forests must proceed. Managers too often forget that they are managing islands of relatively undeveloped habitat in a sea of disturbance. They need to be reminded.

Finally, surveys of biodiversity can also be used in conjunction with other resource inventories and assessments. In particular, inventory personnel can and should seek to highlight cases where compatible land-use demands such as hunting, fishing, and other recreation activities, *support* the need to protect areas for biological diversity.

CONCLUSION

In closing, the Forest Service has clearly made a strong and important commitment to the design and implementation of broad-based timber surveys. As a result, perhaps no country in the world knows as much as ours about the growth and productivity of its timber resources. There is now, however, an urgent and compelling need to establish a similar commitment to inventories of biological diversity.

The preservation of diversity is an issue of global importance with strong and growing public and congressional support. Forest Service Inventory and Analysis units are well positioned to avail themselves of this support but are currently hampered by a lack of adequate coverage in their existing inventories and the perception that timber remains a higher priority.

Establishing a separate conservation inventory tightly focused on priority species, habitats, and threats would serve to eliminate this perception and at the same time make a valuable contribution to the preservation of our nation's natural heritage.

REFERENCES

Allendorf, F.W. and R.F. Leary. 1986. Heterozygosity and fitness in natural populations of animals, pp. 57-76. *In* Michael Soulé (ed.) *Conservation biology: The science of scarcity and diversity.* Sunderland, Massachusetts: Sinauer Associates, Inc.

Brooks, R.T. 1989. Wildlife Habitat Assessment in the Northeastern United States. In *Proceedings of State of the Art Methodology of Forest Inventory.* Syracuse, New York, July 30.

Craver, G.C. 1982. *Multiresource Inventories — A Technique for Determining the Distribution and Extent of Honeysuckle on Commercial Forest Land in South Carolina.* USDA–Forest Service Research Note SE-317. 11 p.

Crow, T.R. 1988. Biological diversity: Why is it important to foresters? *In* J.E. Johnson (ed.) *Managing north central forests for nontimber values.* Washington, D.C.: Society of American Foresters.

Davis, M.B. and C. Zabinski. In press. Changes in geographical range resulting from greenhouse warming. In *Consequences of Greenhouse Warming to Biodiversity*, R. Peters and T. Lovejoy (eds.).

Erwin, T.L. 1982. Tropical forests: Their richness in Coleoptera and other Arthropod species. *The Coleopteristis Bulletin.* 36(1):74-75.

Flather, C.H. 1986. Wildlife Abundance and Occurrence Models: Application in Regional Resource Planning. In *Perspectives on Land Modeling.* Workshop proceedings. Polyscience Publications, Inc.:Montreal, Canada.

Harris, L.D. 1984. *The fragmented forest: Island biogeography theory and the preservation of biotic diversity.* Chicago, Illinois: The University of Chicago Press. 211 p.

Janzen, D.H. 1986. The eternal external threat, pp. 286-303. *In* Michael Soulé (ed.) *Conservation biology: The science of scarcity and diversity.* Sunderland, Massachusetts: Sinauer Associates, Inc.

Ledig, F.T. 1986. Heterozygosity, heterosis, and fitness in outbreeding plants, pp. 77-104. *In* Michael Soulé (ed.) *Conservation biology: The science of scarcity and diversity.* Sunderland, Massachusetts: Sinauer Associates, Inc.

Mintzer, I.M. 1987. *A matter of degrees: The potential for controlling the greenhouse effect.* Washington, D.C.: The World Resources Institute. 60 p.

Norse, E.A., K.L. Rosenbaum, D.S. Wilcove, B.A. Wilcox, W.H. Romme, D.W. Johnston, and M.L. Stout. 1986. *Conserving Biological Diversity in our National Forests.* Washington, D.C.: The Ecological Society of America for The Wilderness Society. 116 p.

Ohman, J.L. 1989. Considering Wildlife Habitat Resources in Forest Inventory and Analysis in the Pacific Coast States — Progress and Future Direction. In *Proceedings of State of the Art Methodology of Forest Inventory*. Syracuse, New York, July 30.

Ralls, K., P.H. Harvey, and A.M. Lyles. 1986. Inbreeding in natural populations of birds and mammals, pp. 35-56. *In* Michael Soulé (ed.) *Conservation biology: The science of scarcity and diversity.* Sunderland, Massachusetts: Sinauer Associates, Inc.

Rudis, V.A. 1988. *Nontimber Values of East Texas Timberland.* USDA–Forest Service, Resource Bulletin, SO-139. New Orleans, Louisiana.

Solheim, S.L., W.S. Alverson, and D.M. Waller. 1987. Maintaining biotic diversity in national forests: The necessity for large blocks of mature forests. *Endangered Species Technical Bulletin* 4(8):1-3.

USDA–Forest Service 1986. *Final environmental impact statement, land and resource management plan, Cherokee National Forest.*

USDA–Forest Service 1988a. *Land Areas of the National Forest System as of September 30, 1988.* FS-383. 87 p.

USDA–Forest Service 1988b. *National Forest System Threatened, Endangered, and Sensitive Species Program*. 15 p.

Wilcove, D.S. 1986. Logging versus biological diversity in the Greater Yellowstone Ecosystem. *Public Lands*. December:10-11.

Wilcove, D.S. 1987. From fragmentation to extinction. *Natural Areas Journal* 7(1):23-29.

Wilcove, D.S. 1988. National forests: Policies for the future. Vol. 2. *Protecting Biological Diversity.* Washington, D.C.: The Wilderness Society. 50 p.

Winjum, J.K. and R.P. Neilson. 1989. Projecting the impacts of greenhouse warming: Forests. *The EPA Journal* 15(1):11-13.

The Forest Inventory:
Incorporating Fisheries and Wildlife Values

Frances A. Hunt
Forestry Resource Specialist
National Wildlife Federation
Washington, D.C. 20036

ABSTRACT

Continued Forest Service effort and enthusiasm are necessary to seek new ways to address nontimber (fish and wildlife, recreation, aesthetics, etc.) values through the inventory process. Collections of relative information must be conducted in a process that balances the realities of regional resource variability and the need for consistent data bases and data products nationally. Improved coordination and communication between forest managers and inventory specialists, as well as improved cooperation with resource scientists will also be required. Additional information regarding nontimber vegetation, dead and downed material, canopy conditions, water resources, and landscape characteristics will be necessary to address fish and wildlife concerns.

INTRODUCTION

Historically, forest inventories have focused on measuring data that are directly relevant to the growth and yield of timber on our National Forests and private lands. These inventories have gathered information to be used in crafting silvicultural prescriptions for forest stands -- prescriptions that have emphasized the production of lumber and fiber from these lands.

Recently, however, greater emphasis has been placed on the evaluation and the management of a wider range of forest resources. Our national policy of multiple use management has created a need for multi-resource inventories to provide the information necessary to avoid and resolve resource conflicts. It is no longer sufficient to determine the value of a particular forest stand for timber production. Although this paper will focus on the inventory needs of wildlife and fisheries resources, the agency must protect and produce a wide range of other non-timber values as well, including: soils, recreation, water resources, and range. Moreover, the agency must address the management of particular stands in light of the ecological requirements and overall management goals for all forest and rangeland resources on a local, regional and national basis.

PROGRESS AND FUTURE CHALLENGES

Already much work has been done to incorporate additional nontimber measures into data collection protocols (Ohmann 1983, Hoekstra et al., 1987). Recently, a co-worker of mine at the National Wildlife Federation, Rudy Rosen, visited the forest inventory unit at Starkville, MS. He remarked on the unit's "strong sense of purpose" and noted that its "scientists actively seek new ways to use their data to address questions about wildlife, fish, and recreation on forests." This level of effort and enthusiasm -- which can be found in many parts of the agency's inventory program -- has resulted in significant progress in broadening the scope of the forest inventory process.

This effort and enthusiasm must continue to build if the challenge to addressing nontimber (wildlife, recreation, aesthetics, etc.) values is to be met. A growing national population, with increased leisure time, will guarantee that resource use and its related conflicts will escalate. Consequently, there will be an expanding need for better resource information to help resolve complex resource conflicts in the years to come.

Another challenge to the usefulness of the forest inventory and the information that it provides will come from within the Forest Service itself. Improved management and information technologies in other aspects of the agency will drive the search for improvements to both the nature and the quality of inventory data. Both GIS and the corporate date base for resource information will challenge those associated with the forest inventory to generate ever more useful and consistent data in the future. In the face of these changes, one of the greatest opportunities will be the crafting of a forest inventory process that provides the consistency necessary to satisfy the requirements of national data bases, while still providing the flexibility necessary to meet information requirements on a regional or local level.

Speaking from the perspective of the National Wildlife Federation, I believe the Forest Service must continue its work to broaden the scope of the inventory to provide more complete information on fisheries and wildlife resources. Currently, it appears that there is great variability between and even within regions in the degree to which appropriate fisheries and wildlife information is gathered during the inventory process. Information must be gathered -- in all regions -- to allow the classification and assessment of wildlife habitats in our nation's forests. This information will be necessary if the Forest Service is going to make informed, professional judgments concerning the implications and trade-offs associated with multiple use management.

Additional improvements to the forest inventory will require improved communication between resource managers, inventory scientists, and other users of inventory data -- both within the agency and outside of it. Improving the inventory can thus be seen as something of a two step -- albeit somewhat circular -- process, which may have to overcome administrative barriers separating agency land managers and inventory specialists. First, communication and cooperation between ecologists, fisheries biologists, wildlife managers, and inventory scientists must improve. Second, we must work with other nontimber interests to broaden the application of the data collected. It has been suggested that certain scientists and resource managers at appropriate levels be assigned the responsibility for insuring that this cooperation actually occurs. Perhaps, for example, FIA analysts should communicate more directly with National Forest System decision-makers. The information gathered through improved communication and cooperation will allow the agency to continue to develop a forest inventory that will provide the information necessary to assess the type, status, and location and range of habitat types found in our public and private forestlands.

MINIMUM FOREST ATTRIBUTES

Given the controversies and conflicting scientific interpretations that seem all but inevitable on our National Forests in the future, there is considerable advantage to an inventory process that identifies and measures specific forest attributes. The collection of data concerning fundamental attributes of the forest allows the subsequent classification of these attributes in a manner that could thus be considered "definition-neutral."

Although it is somewhat difficult to provide specifics to a national forum, I would recommend that -- at a minimum -- the forest inventory conducted in each region of the country should collect information regarding the following five forest attributes:

1. Nontimber Vegetation. While conventional forest inventories have provided a wealth of information about commercial timber, they have not sufficiently or uniformly addressed the so-called "lesser vegetation" that plays such a key role in determining the type and quality of available wildlife habitat. Although some regions have done quite a bit of work in this area, additional work is still needed on linkages between measurements of lesser vegetation and habitat suitability. Forest inventory scientists in each region therefore face an exciting challenge: to cooperate with their ecologist and wildlife co-workers to identify those plant species (including grasses, herbs, and shrubs) whose

presence, quantity, and condition determine and characterize the wildlife habitat being inventoried.

Collection of such additional data (perhaps including total vegetation surveys) can provide estimates of "habitat value," which can be of benefit in making assessments regarding wildlife management. This could perhaps be accomplished on the basis of a stratified random design that allowed an in-depth look at only certain forest types or only in certain geographic areas.

2. Dead and Down Material. In addition to including expanded information on a wider range of non-tree species, additional effort needs to be focused on standing and downed dead material. Although dead, hollow, rotten, or rough trees may have little or now value as a source of fiber or lumber, these same trees are extremely valuable for wildlife habitat. Although the exact manner in which wildlife will utilize these dead and down trees may vary from region to region -- and therefore the exact measurements that must be taken to correctly address habitat needs may vary -- it is vitally important that greater emphasis needs to be placed on the measurement of this resource. In fact, such habitat is now the focus of a major Forest Service initiative to identify the value and prevent the destruction of these trees. Regionally appropriate measures of this forest characteristic need to be collected, nationwide.

3. Canopy Conditions. Another area that may warrant greater attention is the structure of the forest canopy and the degree of closure. A variety of bird species, for example, utilize different canopy layers to varying degrees within a forest (Sheffield 1981). The occurrence and composition of these canopy layers can greatly influence avian populations within the forest.

4. Water Resources: Finally, I would recommend that the inventory gather additional information regarding water resources. The presence and nature of wetlands, the type and quality of streams and other flowing or standing water resources play a significant role in providing and determining the characteristics and quality of available fish and wildlife habitat.

5. Landscape Characteristics: The distribution of various wildlife species are not just dependent upon the vegetative characteristics at a given inventory point. Wildlife populations are also influenced by the size, character, and distribution of surrounding forest stands. The non-forest context of the area is also important, as cropland, pastures, urban areas, all have a significant impact on wildlife resources. These important influences should be addressed in the inventory.

The uniform collection of information concerning these five attributes will greatly improve the applicability of forest inventory data to the management of fish and wildlife resources. The achievement of these improvements will require the continued leadership of researchers and units in every region of the country.

REFERENCES

Ohmann, Janet., 1983. Evaluating Wildlife Habitat As Part of a Continuing, Extensive Forest Inventory. In: Proceedings of the International Conference, Renewable Resource Inventories for Monitoring Changes and Trends; 1983 August 15-19; Corvallis, OR. O.S.U.; p. 623-627.

Hoekstra, Thomas W., et al., 1987. Regional Fish and Wildlife Habitat Models: Application for National Multiple Resource Planning. Paper presented at the Land and Resource Evaluation for National Planning in the Tropics International Conference and Workshop; 1987 January 25-31; Chetumal, Mexico.

Sheffield, Raymond M., 1981. Multiresource Inventories: Techniques for Evaluating Nongame Bird Habitat. Forest Service Research paper SE-218. United States Department of Agriculture. Southeastern Forest Experiment Station. Asheville, NC.

Linking multiple resources analyses to land use and timber management: application and error considerations.

Curtis H. Flather
Research Wildlife Biologist
USDA, Forest Service
Rocky Mountain Forest and Range Experiment Station
240 West Prospect
Fort Collins, Colorado 80526

Linda A. Joyce
Range Scientist
USDA, Forest Service
Rocky Mountain Forest and Range Experiment Station
240 West Prospect
Fort Collins, Colorado 80526

Rudy M. King
Biometrician
USDA, Forest Service
Rocky Mountain Forest and Range Experiment Station
240 West Prospect
Fort Collins, Colorado 80526

ABSTRACT

Natural resource planning conducted by land-managing agencies is required to be multiple resource in context, and regional and national in scope. A modeling framework to address these requirements is reviewed. Results of a case study in the southern United States demonstrate the feasibility of using extant regional land base inventories to assess multiple resource response to timber management and land use trends. Feasibility notwithstanding, an error analysis of resource models is required if tenable and specific recommendations to land base inventory design and analysis are to be made.

INTRODUCTION

Resource management has evolved from qualitative application of professional expertise, to creation of information retrieval systems, to specification of quantitative relationships between environmental factors and resource production. This evolution is not one of replacement, but a continual development of complementary approaches that provide resource planners with an improved basis from which to make informed resource management decisions.

The motivation for developing quantitative analysis techniques has, in part, come from legislative mandates requiring land managing agencies to conduct comprehensive resource planning. The Renewable Resources Planning Act of 1974 requires the Forest Service to conduct periodic national assessments of all renewable natural resources. This law has two important implications to the required national assessment of natural resources. First, the assessment must be completed in a multiple resource context—evaluating resource status must acknowledge competing resource uses. Second, the scope of assessments necessitates developing analyses that are appropriate for national scale evaluations.

The objective of this paper is to review these implications in a case study of multiple resource impacts associated with land use and timber management. A description of the modeling approach is followed by the results of an application evaluating multiple resource production across the South. A discussion of error sources then serves as a focus for concluding remarks concerning land base inventory design and evaluation.

MODELING APPROACH

The Multiple Resource Framework

A major problem in developing multiple resource analyses has been defining a conceptual basis for linking resource areas. Joyce et al. (1986) proposed an analysis framework that integrates regional resource models for use in large-scale assessments of natural resources. The framework links models for forage, wildlife, fish, and water, to models that project areal changes to land base composition and was applied to an analysis of the Southern timber situation (USDA, Forest Service 1988). The study area extended over a 12-state region from east Texas and Oklahoma across the Gulf Coast states and north to Virginia.

Future area estimates for crop, pasture and range land, human related land uses (urban, roads, farm structures, etc.), and timber management types (natural pine, planted pine, oak-pine, upland hardwood, and lowland hardwood) were projected for each state by the Southeastern Area Model (SAM) (Alig 1986). Forest stand characteristics, including growing-stock volume, land ownership, timber site class, and age class, were modeled by the timber inventory model (TRIM) (Tedder et al. 1987) and the timber assessment market model (TAMM) (Adams and Haynes 1980).

To link other resource models to land use and timber projections, all models were required to operate from a commonly defined set of land base descriptors (Joyce et al. 1986). Linkage through a commonly defined land base classification was less constraining than requiring all resource models to function at a common geographic scale since it permitted selection of geographic units that were most appropriate for each resource.

Data Sources and Resource Model Development

Regional land base data were derived from Forest Inventory and Analysis (FIA) surveys (USDA Forest Service 1985) and the 1982 National Resources Inventory (NRI)(USDA Soil Conservation Service and Iowa State University Statistical Laboratory 1987). Coupled with forage, water, wildlife, and fish inventories from federal, state, and private agencies, this land and resource data served as the basis from which empirical models were developed.

Forage production was modeled differently for forestland and pasture/rangeland. On forestland, regression models were developed relating forage production to stand characteristics using plot-level FIA data (Joyce 1988). As a percentage of the original standard deviation, the model standard errors were 10% to 40% less variable than the original data. On pasture and range, a fixed production rate was assigned to each cover type which, when multiplied by the area in each type, gave estimates of forage production (Joyce 1988).

The wildlife and fish models translate land base characteristics into habitat suitability based on discriminant functions (Johnson and Wichern 1982) that predict occurrence or density classes of selected species (Flather 1988, Flather et al. 1989, Flebbe et al. 1988). Because wildlife and fish resource response is a function of cumulative land use and land cover changes over a landscape, a relatively large

Table 1. Projected percent composition of the southern land base.

Land Variable	1985	1990	2000	2010	2020	2030
Cropland	16.7	16.8	16.9	16.7	16.9	17.0
Pasture/Range	15.6	15.3	14.8	14.2	13.8	13.4
Human-related	7.3	7.9	8.9	10.4	10.4	11.0
Forest	56.2	55.9	55.3	54.6	54.7	54.4
Natural Pine	12.7	11.8	9.3	8.0	7.6	7.4
Planted Pine	6.4	8.1	11.6	13.3	14.3	14.8
Oak-Pine	8.3	7.7	7.1	6.8	6.7	6.6
Upland Hardwood	19.5	19.0	18.3	17.7	17.6	17.3
Lowland Hardwood	9.4	9.3	9.0	8.7	8.5	8.3
Age[1]Class 1 (0-20 yrs)	13.8	10.5	13.1	15.4	16.7	17.0
Age Class 2 (20-50 yrs)	28.1	28.3	22.2	16.2	13.2	14.9
Age Class 3 (50+ yrs)	8.1	8.0	8.4	9.6	10.4	7.6
Hardwood[2]Age Class 1	9.7	7.3	8.0	11.2	12.6	12.8
Hardwood Age Class 2	20.2	21.2	18.3	13.0	10.1	11.9
Hardwood Age Class 3	7.3	7.5	8.0	9.2	10.2	7.5
Pine[3]Age Class 1	7.1	6.5	8.1	8.5	7.9	7.1
Pine Age Class 2	12.4	11.9	7.5	5.8	5.9	6.6
Pine Age Class 3	1.5	1.1	0.7	0.6	0.5	0.3

[1] includes natural pine, oak-pine, upland hardwood, and lowland hardwood
[2] includes oak-pine, upland hardwood, lowland hardwood
[3] includes natural pine and oak-pine

geographic area was chosen as the observational unit. For wildlife, the county was used; for fish, the watershed was used. Wildlife and fish data existed for white-tailed deer *(Odocoileus virginianus)*, wild turkey *(Meleagris gallopavo)*, red-cockaded woodpecker *(Picoides borealis)*, and trout (family *Salmonidae*). South-wide classification accuracies of the discriminant models were significantly better than a random model ($p < 0.001$, Kappa statistic; Cohen 1968, Titus et al. 1984).

Water yield models were developed from synthesized information from catchment experiments, state-of-the-art papers, personal knowledge (hydrologists from four physiographic provinces in the South), and other references such as climatological atlases (Ursic 1987). This information was used to construct matrix-response models such that each cell in the matrix represented either the water yield for each land type (the diagonal) or the net increase or decrease in yield when land type changes (off-diagonal). Applying these matrices to projected changes in land use resulted in predicted changes in water yields.

APPLICATION RESULTS

Possible impacts on multiple resource production from changes in land use and timber management were evaluated by using SAM and TRIM/TAMM as input into the resource production models. A postulated future for land use and timber management was constructed based on assumptions concerning human population growth, economic growth, and timber management for a planning period from 1985 to

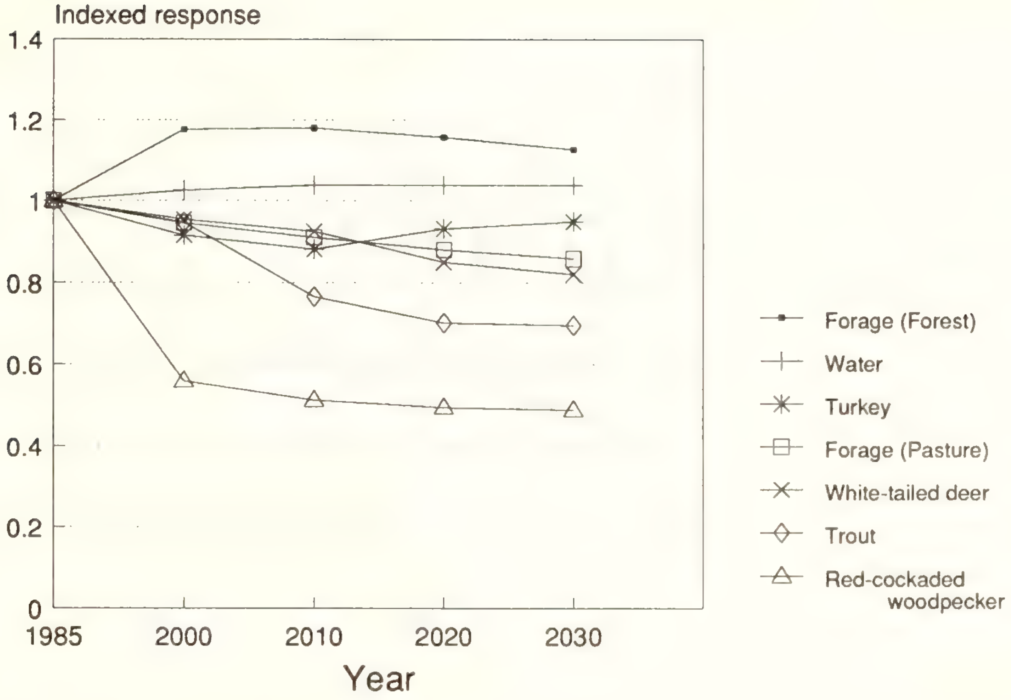


Figure 1. Indexed resource response to changes in land base characteristics from 1985 to 2030. Resource values over time are expressed relative to 1985.

2030 (USDA, Forest Service 1988). Under this assumed future, land area changes are characterized by an overall reduction in forestland and pasture/rangeland, which are converted to cropland and human-related land uses (table 1). Acres in planted pine are projected to increase from about 6.5% of the land base to nearly 15%. The conversion of natural pine stands to pine plantations accounts for the majority of the increase. Projected changes in forest stand age were characterized by increases in both younger and older age classes of hardwood types, and declines in older age classes of pine.

Region-Wide Summaries of Resource Response

Projected resource responses to these land base trends vary (fig. 1). Water and forage production from forestland are the only resources that increase over the projection period. Increases in forage reflect increases in pine types and younger age classes of timber as older stands are harvested and regenerated. The slight increase in water yield is primarily a response to increased land area in human-related land uses.

Trout, deer, and turkey all show moderate declines. Trout decline in response to the harvesting of older aged hardwoods in the coldwater region of the South. Turkey density also declines initially but recovers during the later periods as a large acreage of moderate-aged hardwoods mature in the Southcentral. Deer decline throughout the projection period in response to loss of natural timber types and the significant increase in human-related land use.

Red-cockaded woodpeckers show the greatest relative decline of all resources modeled. Dependance on mature pine stands results in a notable loss of active nesting colonies as natural pine types are brought under plantation, short-rotation management.

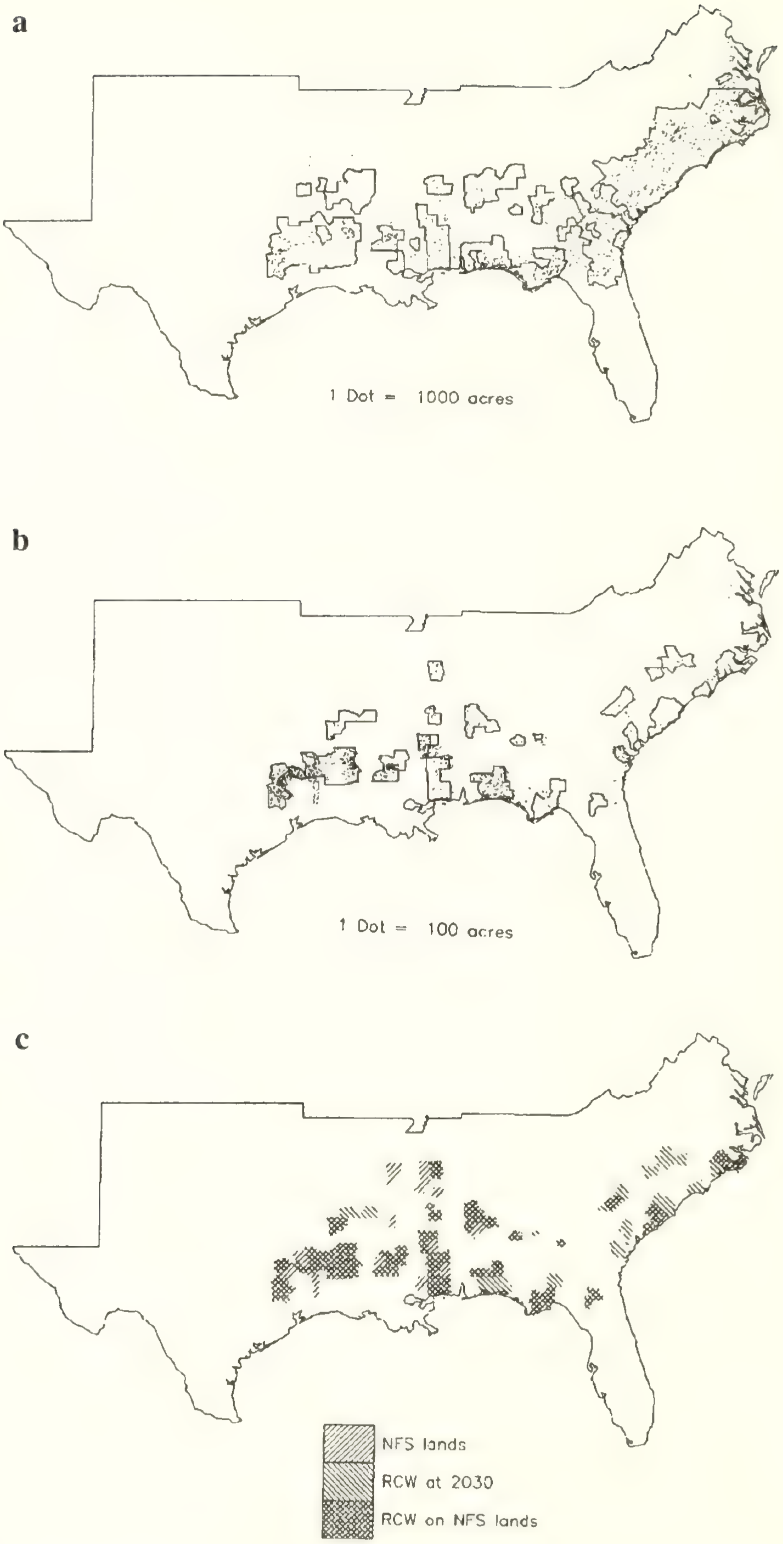


Figure 2. Distribution of counties supporting active red-cockaded woodpecker nesting colonies (counties with active colonies are outlined) in relation to the distribution of older age classes of natural pine (as represented by dots) in 1985 (a) compared to the predicted distribution in 2030 (b). The predicted distribution of active colonies in 2030 by county compared to the distribution of counties with national forest ownership (c).

Spatial Aspects of Resource Response

Using the red-cockaded woodpecker as an example, additional insights into land-base relationships and regional resource production can be gained from examining spatial patterns across the region. The close tie between red-cockaded woodpecker nesting sites and mature pine stands is obvious when the distribution of old pine in the primary range of the bird is overlayed with those counties currently supporting active colonies (fig. 2a). This relationship is maintained over the projection period. Counties predicted to retain active colonies in 2030 are also those projected to retain some acreage of mature pine (fig. 2b). Such spatial displays provide more detailed descriptions of resource response and are also of heuristic value. One can ask if there are other land base associations that help explain the distribution of active nesting sites at the end of the projection period. Land ownership is one possible factor. Private timberland was managed more intensively than public timberland in TRIM resulting in a strong correspondence between retention of suitable red-cockaded woodpecker nesting habitat and national forests (fig. 2c). This particular projection indicates that, unless private timber management considers the bird's habitat requirements, public lands will be critically important to the preservation of this threatened and endangered species.

SOURCES AND IMPLICATION OF ESTIMATION ERRORS

The preceding discussion has acknowledged neither the existence of error nor the implications of that error on output interpretation. Although standard "goodness of fit" measures indicate that the developed resource models fit the data reasonably well, there is a need for further research that rigorously examines sources of error and how they affect inventory and modeling aspects of the multiple resource framework applied here.

Estimation error is affected by many components of the modeling process: precision, accuracy, and representativeness of input data for model estimation, model specification in an efficient mathematical form, propagation of error through time and from one model to others, and aggregation of model output to alternative geographic scales. Examining these various sources of error with respect to the multiple resource modeling framework illustrates the complexity and potential magnitude of the problem.

If a resource, Y, is modeled as responding to land base factors, $X_1, X_2, ..., X_k$, the relationship can be represented by

$$Y = f(X_1, X_2, ..., X_k; b_1, b_2, ..., b_p) + \varepsilon \qquad [1]$$

where f(.) represents the functional form of the model, b_i represents one of p model coefficients, and ε represents unmodeled residual variation. A general equation estimating the error or variance for an estimate of Y is given by (Kempthorne and Folks 1971)

$$\hat{Var}(\hat{Y}) = \sum_{m}^{p}\sum_{n}^{p} (\partial\hat{Y}/\partial b_m)(\partial\hat{Y}/\partial b_n) S_{b_m} S_{b_n} R_{b_m b_n} + \sum_{i}^{k}\sum_{j}^{k} (\partial\hat{Y}/\partial X_i)(\partial\hat{Y}/\partial X_j) S_{x_i} S_{x_j} R_{x_i x_j} \qquad [2]$$

where $\hat{Y}$ is an estimated value of Y for specified values of the factors X_i, and S_{x_i} and S_{b_m} are the standard deviations of X_i and b_m, respectively. $R_{b_m b_n}$ and $R_{x_i x_j}$ are pairwise correlations between model coefficients and factors, respectively. Additional sums could be included in [2] to account for correlations between factors and model coefficients, but are eliminated here for simplicity.

The structure, complexity, and size of the estimation error depends on attributes of information used to specify model components. For a single estimation of an individual Y, if specified values of the factors are known without error, then only terms from line one of [2] are needed to estimate variance. Further, if factors are uncorrelated and model structure is such that model coefficients are uncorrelated, then only a simple sum of model coefficient variances is needed to estimate variance (i.e., only the m=n terms in the first line of [2]). If factors or model coefficients are correlated, then estimated variance is increased by the m≠n terms in the first line of [2].

If factor information is only estimated (e.g., when sample estimates are used to initialize a model computation, when a model is used for multiple time projections, or when one model is used to initialize another model), then the i=j terms from the second line of [2] are added to the variance computation. If correlations exist among factors, then the i≠j terms in the second line of [2] are also included in the variance computation. For multiple time projection models or models linked in series, variance attained at any step feeds directly into succeeding variance computations through the factor terms, increasing at each step.

By now it should be clear that error estimation of model estimates can be very complex, often including variance and correlation information that is difficult to obtain. However, blind acceptance of model outputs is a poor alternative, because estimation error can be substantial. Obviously, it is desirable to control the error associated with as many of these components as possible. Further, it is desirable to estimate the error associated with as many of these components as possible so that realistic error estimates can be attached to model output, thus establishing a perspective for incorporating model estimates into decision making processes. Only then can we estimate, for example, whether projected water yield "increases" indicated in Figure 2 should be expected.

RECOMMENDATIONS AND CONCLUSIONS

The multiresource framework was designed to evaluate the potential effects of land use shifts and timber management on the forage, water, wildlife, and fish resources. This framework, a first-generation approach to modeling multiple resource response at the regional level, has been successfully applied in analyzing the southern timber resource management policy (USDA, Forest Service 1988). Although this study has demonstrated feasibility of using extant land base inventories to model regional resource responses, there is a need to quantify model performance based on a rigorous examination of error. Quantification of error is a critical step in improving both resource models and, more basically, the inventories that support empirical analyses. The latter is particularly important given the design of Forest Service multiple resource inventories. Data elements important for resources such as wildlife or forage have been appended to an inventory designed to measure timber characteristics within desired precision limits. There is no guarantee that the error associated with measuring land base attributes for other resources will fall within acceptable limits. Only by quantifying the many sources of error will multiple resource inventories and the models that these inventories support be improved and the uncertainty associated with multiple resource planning decisions be explicit.

REFERENCES

Adams, D. M., and R. W. Haynes. 1980: Softwood timber assessment market model: Structure, projections, and policy simulations. Forest Science Monograph no. 22, 64 pp.

Alig, R. J. 1986: Econometric analysis of forest acreage trends in the Southeast. Forest Science 32:119-134.

Cohen, J. 1968: Weighted kappa: nominal scale agreement with provisions for scaled disagreement or partial credit. Psychological Bulletin 70:213-220.

Flather, C. H., T. W. Hoekstra, D. E. Chalk, N. D. Cost, and V. A. Rudis. 1989: Recent historical and projected regional trends of white-tailed deer and wild turkey in the southern United States. U.S. Department of Agriculture, Forest Service, General Technical Report RM-172. Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO. 22 pp.

Flather, C. H. 1988: Wildlife abundance and occurrence models: application in regional resource planning. Pages 37-47 in R. Gelinas, D. Bond,and B. Smit, (eds.) Perspectives on Land Modelling. Workshop Proceedings. 1986 17-20 November, Toronto, Canada. Polyscience Publications, Inc.

Flebbe, P. A., N. Cost, and T. W. Hoekstra. 1988: Historical and Projected Trout Populations in the southern United States. U.S. Department of Agriculture Forest Service General Technical Report RM-160. Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO, 19 pp.

Johnson, R. A. and D. W. Wichern. 1982: Applied multivariate analysis. Prentice-Hall, Englewood Cliffs, NJ, 594 pp.

Joyce, L. A. 1988: Regional forage model. Pages 17-26 in R. Gelinas, D. Bond, and B. Smit, (eds.) Perspectives on Land Modelling. Workshop Proceedings, 1986 17-20 November, Toronto, Canada. Polyscience Publications Inc.

Joyce, L. A., T. W. Hoekstra, and R. J. Alig, 1986: Regional multiresource models in a national framework. Environmental Management 20:761-771.

Kempthorne, O. and L. Folks, 1971: Probability, statistics, and data analysis. Iowa State University Press. Ames, IA, 555 pp.

Tedder, P. L., Richard N. LaMont, and Jonna C. Kincaid, 1987: The timber resource inventory model (TRIM): a projection model for timber supply and policy analysis. U.S. Department of Agriculture Forest Service General Technical Report PNW-202. Pacific Northwest Research Station, Portland, OR., 82 pp.

Titus, K., J. A. Mosher, and B. K. Williams, 1984: Chance-corrected classification for use in discriminant analysis: ecological applications. The American Midland Naturalist 11:1-5.

Ursic, S., 1987: The South's fourth forest: regional water response to timber management. Supplementary Analyses. Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO, 22 pp.

U.S. Department of Agriculture, Forest Service, 1988: The South's Fourth Forest. Forest Resource Report No. 24. Forest Service, Washington, DC., 512 pp.

U.S. Department of Agriculture, Forest Service, 1985: Forest Service resource inventory: an overview. Forest Resources Economics Research Staff. U.S. Department of Agriculture, Forest Service, Washington, DC. 29 pp.

U.S. Department of Agriculture, Soil Conservation Service and Iowa State University Statistical Laboratory, 1987: Basic statistics 1982 national resources inventory. Statistical Bulletin 756. U.S. Department of Agriculture, Soil Conservation Service, Washington, DC. 153 pp.

Using Land Base Inventories To Assess Range And Wildlife Resources Nationally

Linda A. Joyce
Range Scientist
USDA, Forest Service, Rocky Mountain Forest and Range Experiment Station
Fort Collins, Colorado 80526

Curtis H. Flather
Research Wildlife Biologist
USDA, Forest Service, Rocky Mountain Forest and Range Experiment Station
Fort Collins, Colorado 80526

ABSTRACT

The 1989 Assessments of range, wildlife, and fish resources have been greatly facilitated by the continual improvements in land inventories. Recommendations are made concerning future inventories to further improve the supply and demand analyses of range and wildlife resources.

INTRODUCTION

The Renewable Resource Planning Act of 1974 (RPA, Public Law 93-378 (88 Stat. 475, as amended), directed the Secretary of Agriculture to prepare a decadal assessment that analyzes "the present and anticipated uses, demand for, and supply of the renewable resources of forest, range, and other associated lands with consideration of the international resource situation, and emphasis of pertinent supply, demand, and price relationship trends." The third assessment, completed in 1989, represents an increasing dependence on land inventories and analytical tools to describe the resource/land connection.

Our objective in this paper is to describe the use of inventory data in the most recent national assessment as they were used to summarize current conditions and analyze future trends for wildlife, fish, and range resources (Flather and Hoekstra [In press], Joyce [In press]). An additional objective is to show how such analyses might be improved.

CURRENT DESCRIPTION OF RESOURCE PRODUCTION

The data available to support the 1989 assessment of range, wildlife, and fish come largely from existing Forest Service information and cooperating state and federal agencies. In general, the data were not collected specifically for a national assessment. No standard national or regional inventory permits a consistent summarization of wildlife and fish resources (Flather and Hoekstra [In press]) or range resources (Joyce [In press]). Consequently, the extent to which trends in the supply/demand of these resources can be discussed is a function of the information available from various sources.

Range Resource

Range has traditionally been perceived as livestock. However, forage is produced on forest and rangelands, and the 1989 Assessment focused on the supply/demand of range forage. In general, two aspects are important in characterizing the supply and demand of range forage: land availability for forage production, and forest and rangeland productivity. Land availability includes the determination of forest and rangelands that are capable of forage production, and the amount and extent that these lands are grazed. Grazing implies use by large herbivores, wild or domestic. Assessing forage productivity involves a measurement of the quantity currently produced, the quality of forage or species composition, and the potential forage production.

Forest and Rangeland Area for Forage Production. Forage from range and forest land is used in a feed mix involving forage from other sources, such as pastureland and cropland. Assessing the implications of this interconnectedness in forage supply/demand necessitates a consistent description of the historical and current land use. Various agencies inventory land, but no single agency inventories all lands. A historical series of land uses for the entire United States has been summarized by Frey and Hexem (1985). This series represents a compilation of land use data reported by public agencies such as Forest Service, Bureau of Census, Bureau of Land Management, and Soil Conservation Service, and provides a useful framework within which historical and future shifts in the use of the entire land base of the United States can be analyzed (figure 1). The terminology used in this aggregation of inventories reflects an attempt to be consistent over time rather than with one inventory. Frey and Hexem (1985) categorized the major land uses as forest land, cropland, and pasture and rangeland, and the special-use areas as parks, wilderness areas, roads, and industrial areas. Trends in this data reflect use, not land cover.

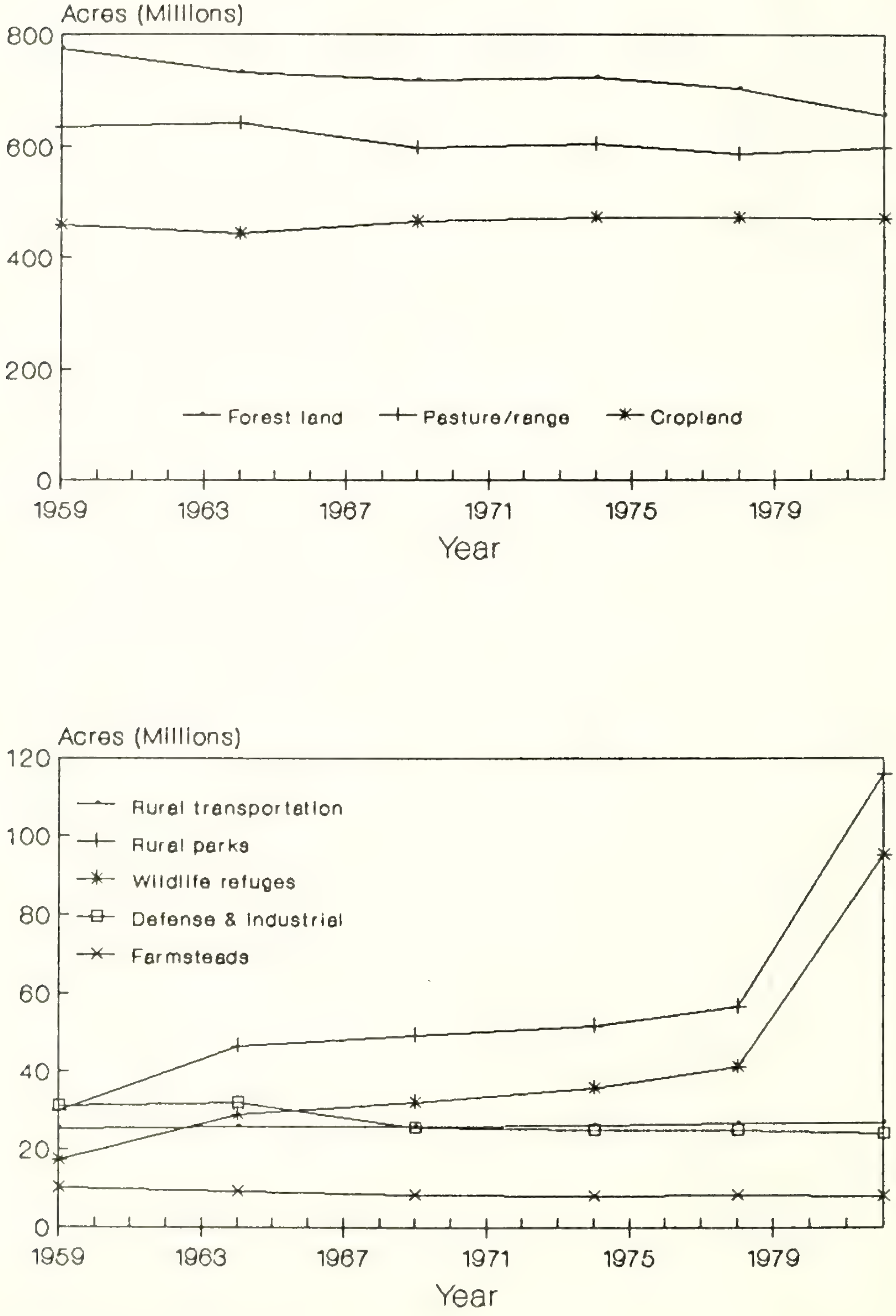
While this data base offers a valuable comparison in land use trends, it is not without problems. Difficulties in classifying forest versus rangeland can influence an interpretation of long-term trends in land use. For example, pasture/rangeland area in the South increased from 1979 to 1982, primarily the result of a change in the definition of rangeland to include some shrub types. These and other assumptions reflect the judgement of the few individuals who have worked on this data, primarily Wooten and Frey. Clearly, a quantitative algorithm linking historical inventories across all agencies/lands would allow the reinvestigation of revisions in the time series. Such an algorithm is not available, and this qualitatively combined data base offers the only consistent look at land use trends at the national level.

The availability of land for grazing, as assessed in the Frey and Hexem series, does not capture whether the land is actually grazed or the degree of use. For example, grazed forest land was estimated from the area in open forest, arid woodlands, and lands reverting to forest which have forage. All pasture and rangeland was assumed to be available for grazing. Land placed in wilderness was assumed to be no longer available for grazing. Frey and Hexem (1985) reported that between 1969 and 1982, grazingland declined 8%. Certainly, a long-term decline in grazingland is the result of forces such as the demand for crop products, the withdrawal of land for recreational, wildlife, and environmental purposes, and for urban areas (Frey and Hexem 1985). These estimates in the long-term series, however, may actually underestimate grazingland as refuges, wilderness, and parks are grazed by livestock. Further, these estimates do not assess grazing by wildlife. Grazing surveys on some ownerships indicate that the extent of lands grazed can change from state to state and year to year. The Soil Conservation Service estimates that from less than 2% to over 75% of pastureland within a state may remain ungrazed annually (U.S. Department of Agriculture and Iowa State University Statistical Laboratory 1987). Rudis et al. (1984) reported that only 8% of the total forest land in Alabama was grazed. Methodologies for these grazing surveys differ. Thus, while these surveys offer important information on the regional extent of grazing, the information is inconsistent at the national level.

The degree of use has been considered important in determining the condition of grazinglands, and has been used to estimate forage production. Grazing use and degree of use are difficult to determine, because use must be separated from regrowth—a problem within a growing season. On rangelands, condition has been traditionally used to determine the degree of use. This term, however, is difficult to interpret (Box 1988).

Forage Productivity. In assessing forage productivity, three characteristics are important: quantity of production, species composition, and current production relative to the site's potential production.

Figure 1--Major uses of land in the United States: 1959-1982.



Source: Frey and Hexem (1985)
Note: Data from periodic inventories.
Most recent inventory: 1982.

At this point in time, consistent estimates of these characteristics are not inventoried nationally. Surrogates such as forage consumption estimates based on livestock inventories and purchased feed supplies have been used to estimate forage production nationally. However, use, not production, is quantified in this method. Forage produced but not consumed is omitted from this estimate, as is forage consumed by uncensused wild grazing animals. Thus, current productivity was described in the Range Assessment using range condition and ecological status (Joyce [In press]). The difficulty of using these surrogates for range productivity is that range condition has been traditionally tied to livestock production, and neither condition nor ecological status are good correlates for forage production (Box 1988). Understory vegetation has been assessed regionally using FIA data (Flather and Joyce in this proceedings), but, the variety of understory measures among regional inventories precludes a national summary.

Wildlife and Fish Resource

In general, there are four aspects of wildlife and fish resources that are important to a characterization of the status of these resources: habitats, population levels, number of users, and harvest levels. The reader is urged to review Flather and Hoekstra [In press] for a full description of the trends in these attributes. The pertinent point for this discussion is that inventories of these aspects of wildlife and fish resources are not necessarily the responsibility of the Forest Service. Thus, coordination with other federal agencies will be necessary to mesh population estimates with the habitat that supports these populations. This paper focuses on the assessment of forest habitat.

Land Uses as a Framework for Wildlife Habitat. Wildlife and fish habitat in its most basic sense can be defined as the availability and appropriate mix of food, cover, and water. Except for special cases (e.g., critical habitat for threatened and endangered species), national inventories assessing the amount of habitat specific to a single species or species group do not exist. Vegetation features can be used to define habitat types that are inventoried over large geographic areas. Based on this definition of habitat, the inventory represents a description and estimate of land area that may support a faunal community as opposed to an areal estimate of suitable habitat for any given species. This alternative definition of the term wildlife and fish habitat forms the basis for the discussion of habitat trends in the assessment.

The relatively minor changes in major land uses as described by Frey and Hexem (1985) indicated the need for a closer examination of shifts within these land uses, particularly forest and rangeland. The variety of forest types, successional stages, and their arrangement across a landscape are factors affecting the numbers and kinds of species inhabiting forest environments.

Measures of Forest Habitat Composition. A detailed analysis for forest land changes was made using FIA data compiled in the timber assessments (USDA Forest Service 1982, Haynes [In press]). The extent of forest land has diminished 4% since 1963 (Flather and Hoekstra [In press]). This loss is less than half that reported by Frey and Hexem (1985), because forest land converted to wilderness does not move out of the FIA forest land category. Forest land loss was greatest in the eastern half of the country, particularly in the South, where forest acres declined by nearly 20 million in the past decade. Forest land acres in the Rocky Mountain and Pacific Coast have remained relatively stable since the early 1960's.

Three attributes of forest land are important to an evaluation of wildlife habitat for national assessments. These attributes include areal trends in forest cover types, areal trends in successional stages, and trends in the size, shape, and spatial distribution of forest types and successional stages. Inventories of forest land have not been uniformly designed to evaluate these descriptive factors. The recent historical trends reported in the Wildlife Assessment are based on surrogates available in the inventory information. Specifically, information exists on trends in forest ecosystem types and successional stages (as measured by stand-size class) for timberland (commercial forest land) only.

As forest ecosystems progress through the sequence of seral stages, the faunal community changes in response to the varying vegetation characteristics of each stage. Maintaining the diversity of wildlife species that are potential inhabitants of any forest ecosystem requires that all seral stages be represented. Stand-size classes for commercial timber were available as indicators of seral stage. Stand-size is defined by the predominant size of trees stocking a stand, and included seedling/sapling, poletimber, sawtimber, and non-stocked stands. In 1987, slightly more than half (242 million acres) of the nation's timberland was classified as sawtimber. The number of acres classified as sawtimber appears to have increased between 1963 and 1987 — a trend due primarily to aging eastern forests. Since 1963, northern sawtimber stands have increased by nearly 22 million acres, or 40%. The area of sawtimber stands has remained relatively stable in the West over the same period.

Stand-size class is not the best indicator of the amount of timberland in mature successional stages. Age, although a better indicator of mature or old-growth forests, is also insufficient. Important structural characteristics, including the presence of snags, dead and down woody material in various stages of decay, multi-layer canopies, and patch understory (Harris 1984, Franklin et al., 1981), need to be considered since they can be absent in intensively managed mature forests. Thus, stand-size class is not the best indicator for old-growth forests and their obligate inhabitants.

Similar data for noncommercial forest land, including parks and wilderness, are not available. Nationally, this omission involves 30% of the total forest land and regionally, for 1% to 63% of the total forest land (Bones [In press]). Even within timberland, variations in inventory techniques and standards cautions any interpretation of historical trends in forest ecosystems types (USDA Forest Service 1982). For western forests, changes in standards and the addition of five million acres not in the 1970 land base preclude making meaningful historical interpretations of vegetation changes (USDA Forest Service 1982). In addition, reported losses of timberland do not necessarily reflect conversion of forest to non-forest lands. Designation of forest land as wilderness removes that land from the timberland base.

Measures of Forest Habitat Configuration. A final characteristic of forest habitats, and one that is inadequately addressed in current forest inventories and assessments, is the size, shape, and distribution of forest land, forest types, and successional stages. There is an increasing recognition that the pattern of forest environments across landscapes needs to be considered in wildlife habitat assessments (Noss 1987). Although some species are benefited by increases in the spatial heterogeneity of forest lands, other species appear to require large tracks of homogeneous forest. There is a concern, both in the East (Burgess and Sharpe 1981) and in the West (Harris 1984), that increasing forest fragmentation will eliminate some species as functioning members of certain regional faunal communities. Methods to evaluate the impacts of changing forest type, size-class, and their interspersion and juxtaposition on wildlife is discussed in the next section.

HOW LAND BASE INVENTORIES CAN BE USED TO SUPPORT PROJECTION ANALYSIS

Wildlife Habitat Evaluations

Evaluating the impacts of changing forest type, stand-size class, and their interspersion and juxtaposition on wildlife and fish is difficult since species will be affected differently, depending on their habitat requirements. Quantitative analyses are being developed to permit resource planners to explicitly analyze species response to changes in forest land characteristics. In a case study for the Wildlife Assessment, Flather and Hoekstra [In press] modified models developed by McClure et al. (1979) and Sheffield (1981) to assess the status and trends in commercial forest habitats for gray squirrel, pileated woodpecker, pine warbler, prothonotary warbler, and red-eyed vireo in the five Atlantic Coastal states from Virginia to Florida. Habitat evaluation models of this type permit examination of status and trends in habitat quality, provided that inventory data elements are temporally consistent (table 1). The development of similar models for other species and regions will require further research before future wildlife assessments can have complete information on wildlife habitat of this nature at the national level.

Table 1.—Analysis of status and trend of habitat for five selected species on commercial forestland in the Southeast and South Carolina.

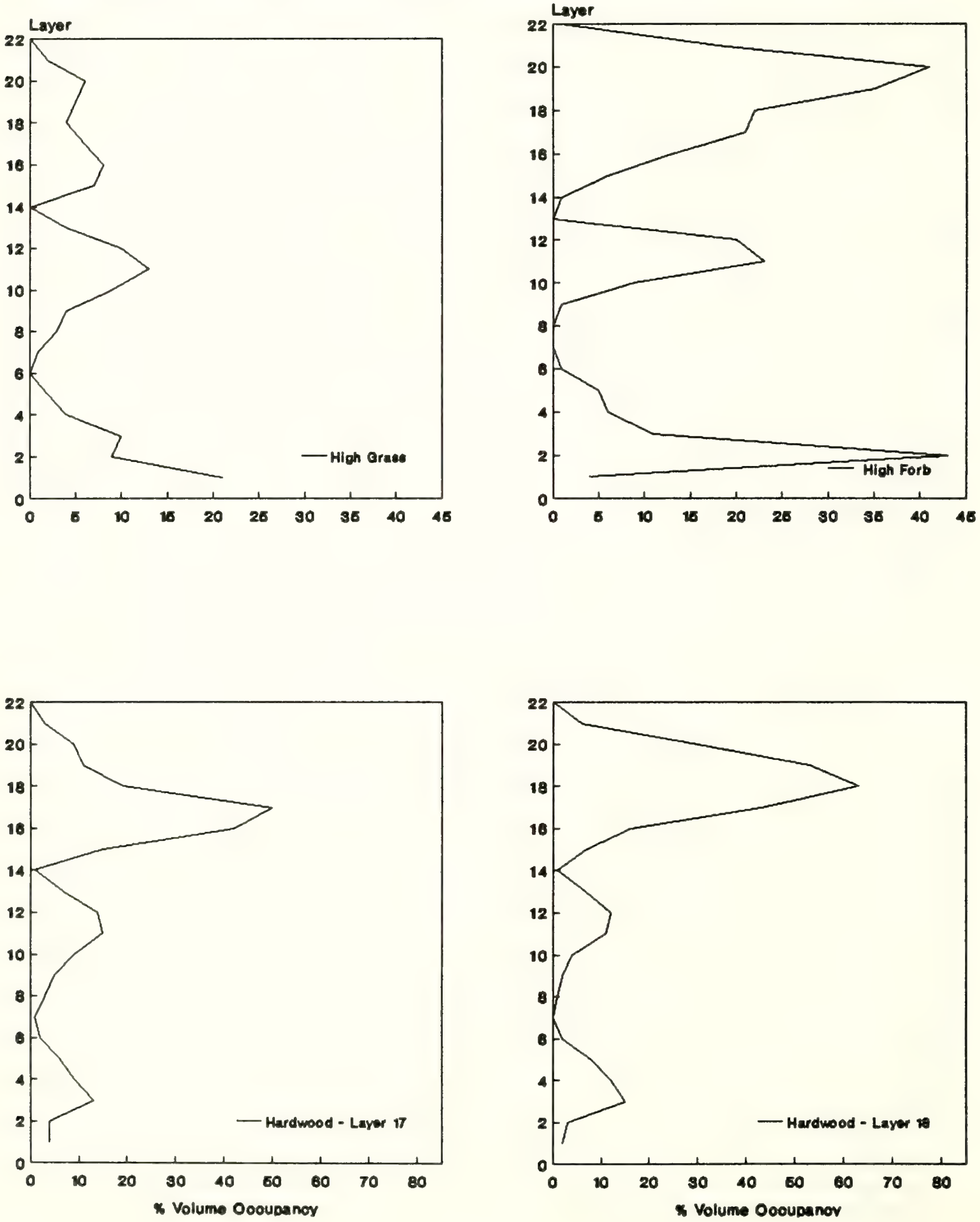
Species	% Good Habitat	% Fair Habitat	% No Habitat
Gray Squirrel			
Southeast	48.5	23.1	28.4
South Carolina 1978	47.4	25.0	27.6
South Carolina 1986	48.5	21.8	29.7
Pileated Woodpecker			
Southeast	7.3	18.5	74.2
South Carolina 1978	7.1	17.7	75.2
South Carolina 1986	6.7	16.3	76.9
Prothonotary Warbler			
Southeast	1.9	2.1	96.0
South Carolina 1978	10.1	6.7	83.2
South Carolina 1986	2.1	2.4	95.5
Pine Warbler			
Southeast	19.5	10.2	70.3
South Carolina 1978	26.9	9.2	63.9
South Carolina 1986	23.8	10.5	65.6
Red-eyed Vireo			
Southeast	18.3	31.1	50.6
South Carolina 1978	9.5	30.1	60.4
South Carolina 1986	14.3	29.6	56.1

Critical questions to be addressed during these and others efforts to link wildlife with land base characteristics are: (1) can existing data sources and analysis methods be used to relate regional patterns of wildlife distribution and abundance to land use and land cover patterns? (2) How reliable are the approaches, and do they offer tenable predictions of wildlife response to changing land management activities? Although such feasibility questions (question 1) are necessary, the probability of achievement is high if evaluation criteria are simply based on successful application. Where data are lacking, assumptions and surrogate measures proliferate which all but assure successful application. Before land planners can use these approaches with confidence, an effort must be made to evaluate how well these efforts perform in light of the assumptions that were made (question 2). We see these efforts to evaluate wildlife habitat models as fruitful areas for collaboration between research wildlife biologists and inventory scientists.

Vegetation Classification for Grazing

Flather and Joyce (this proceedings) described a forage projection model developed from Forest Service inventories of understory biomass and timber stand characteristics. Humphries et al. [unpublished], using FIA data, developed discriminant functions to determine a forest stand's potential for grazing, in the manner that wildlife models describe habitat. Measurements of vegetation volume occupancy at 1-foot intervals from ground to tree canopy top, in some cases over 100 feet, were used to profile the vegetation. These 1-foot measurements from South Carolina timber plots were grouped into 22 vegetation layers, and the plots were clustered into 12 profiles based on volume occupancy similarity across the layers. Plots with high vegetation volume occupancy for grasses separated out from plots with high volume occupancy for forbs (figure 2). Similarly, volume occupancy separated hardwood and softwood plots of increasing volume occupancy through the tree canopy. Thus, sites with good habitat for livestock would be represented by sites with high volume of grasses.

Figure 2. Four profiles of vegetation volume occupancy.



Management variables projected in the timber supply/demand analyses were used to develop discriminant models that would associate projected timber stands with one of the above-defined profiles. Area associated with sites most likely to be grazed could be determined and used as surrogate to project the impact of timber on the forage resource. While timber plots clustered into visibly and statistically different profiles, management variables discriminated sites with distinct overstory canopies (above 6 feet), both hardwood and softwood, better than sites with large amounts of grasses, forbs, and shrubs. Thus, while plots could be assessed with respect to grazing habitat using a temporally consistent inventory, this current description of grazing was not reliably linked to the projection methodology for timber. Vegetation profile may offer an integrative link across all uses of vegetation, such as wildlife habitat and livestock grazing, however, the vegetation profile represents a different approach to assessing forage production than has been traditionally used.

Land Base Projections

The models described above link resource production with forest inventory data, such as forest structure or timber volume. Projecting trends in resource status requires that these land base characteristics be projected into the future. Current land area models focus on a single resource such as timberland, although the ability to simultaneously project land areas for several land uses has been demonstrated for the South (USDA Forest Service 1988). In addition to area estimates by land uses, the pattern of this land use must be distributed across the landscape at scale meaningful to other resources.

RECOMMENDATIONS

The needs of the RPA Assessment are national, and multiresource. Assessment of the current situation of range, wildlife, and fish resources could be improved if:

1. A consistent and complete description of the land base were available for base years;
2. A consistent land description was used to develop national analyses linking resource production to land variables;
3. Inventory scientists and RPA analysts continue and increase their cooperation on research and analysis for resource assessments.

LITERATURE CITED

Bones, J.T., [In press]: An analysis of the lands situation in the United States: 1989-2040. Gen. Tech. Rep. RM-00. U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado.

Box, T. W., 1988: Range Condition. Paper presented at the Resources for the 21 Century. 1988 November. American Forestry Association, Washington, DC.

Burgess, R. L. and Sharpe, D. M., eds., 1981: Forest island dynamics in man-dominated landscapes. Ecological Studies 41. Springer-Verlag, New York, 310 pp.

Flather, C. H. and Hoekstra, T. W., [In press]: An analysis of the wildlife and fish situation in the United States: 1989-2040. Gen. Tech. Rep. RM-00. U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado.

Franklin, J. F., Cromack, K., Denison, W. [and others], 1981: Ecological characteristics of old-growth Douglas-fir forests. Gen. Tech. Rep. PNW-118. U.S. Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station, Portland, Oregon, 48 pp.

Frey, H. T., and Hexem, R. W., 1985: Major use of land in the United States: 1982. Agric. Econ. Rep. No. 535. U.S. Department of Agriculture, Economics Research Service, Natural Resource Economics Division, Washington, DC, 29 pp.

Harris, L. D., 1984: The fragmented forests: Island biogeography theory and the preservation of biotic diversity. University of Chicago Press, Chicago, 211 pp.

Haynes, R. W., [In press]: An analysis of the timber situation in the United States: 1989-2040. Gen. Tech. Rep. RM-00. U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado.

Humphries, H., Joyce, L. A., Cost, N. D. Classifying vegetation characteristics for multiresource production. [in prep.]

Joyce, L. A., [In press]: An analysis of the range forage situation in the United States: 1989-2040. Gen. Tech. Rep. RM-00. U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado.

McClure, J. P., Cost, N. D. and Knight, H. A., 1979: Multiresource inventories—a new concept for forest survey. Res. Pap. SE-191. U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station, Asheville, North Carolina, 68 pp.

Noss, R. F., 1987: From plant communities to landscapes in conservation inventories: a look at the Nature Conservancy (USA). Biological Conservation. vol 41, pp. 11-37.

Rudis, V. A., Rosson, Jr., J. F. and Kelly, J. F., 1984: Forest resources of Alabama. Res. Bull. SO-98. U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station, New Orleans, Louisiana, 55 p.

Sheffield, R. M., 1981: Multiresource inventories: Techniques for evaluation nongame bird habitat. Res. Pap. SE-218. U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station, Asheville, North Carolina, 28 pp.

U.S. Department of Agriculture, Soil Conservation Service and Iowa State University, Statistical Laboratory, 1987: Basic Statistics 1982 National Resources Inventory. Statistical Bulletin No. 756. Washington, DC, 153 pp.

U.S. Department of Agriculture, Forest Service, 1982: An analysis of the timber situation in the United States, 1952-2030. Forest Resource Report No. 23. U.S. Government Printing Office, Washington, DC, 499 pp.

U.S. Department of Agriculture, Forest Service, 1988: The South's fourth forest: alternatives for the future. Forest Resource Rep. No. 24. Washington, DC, 512 pp.

Multiple Value Forest Surveys in the Midsouth States

Victor A. Rudis
Research Forester, Forest Inventory and Analysis Unit
USDA Forest Service, Southern Forest Experiment Station
201 Lincoln Green, Starkville, MS 39759-0906

ABSTRACT

State-of-the-art achievements and limitations in integrating water, range, wildlife, and recreation ("nontimber") inventories with forest surveys of the USDA-Forest Service, Southern Forest Experiment Station, Forest Inventory and Analysis (FIA) Unit are reviewed. The FIA Unit surveys private and public forests in 7 Midsouth states: Alabama, Arkansas, Louisiana, Mississippi, Oklahoma, Tennessee, and Texas. Articles, works-in-progress, and the process of integrating multiple-value forest inventories while maintaining an ongoing forest survey are discussed with respect to: (1) current and projected trends, (2) new approaches and additional monitoring measures, and (3) establishment of a user constituency in nontimber disciplines. Recommendations for the future include studies of sample designs suited to multiresources assessment, focusing analytical reports toward newsworthy issues in nontimber disciplines, fostering interdisciplinary efforts to study regional forest resource issues through cooperative funding programs, and "outreach" activities to inform and involve individuals in the multiple value potential of forest survey information.

INTRODUCTION

The U.S. Department of Agriculture Forest Service (USDA-FS), Southern Forest Experiment Station (SO), Forest Inventory and Analysis (FIA) Unit, has been estimating current conditions and trends in private and public forest resources since its inception in the 1930's. The SO-FIA Unit conducts forest surveys on some 100 million acres of largely private forests in Midsouth states (Alabama, Arkansas, Louisiana, Mississippi, Oklahoma, Tennessee, and Texas). Public forests of the USDA Forest Service's National Forest System are surveyed by the SO-FIA Unit as well for regional analyses.

The forest survey sample design involves field observations of trees in forested areas, taken within 1-acre plots located systematically at 3-mile intervals throughout the Midsouth states. When combined with ground-truth of forest/nonforest photointerpretation on additional areas, field observations of trees and forested areas are expanded statistically to estimate forest area and timber volume for entire counties, states, and regions, regardless of ownership. Such data routinely are compiled, examined, and reported for states and the Nation about every 10 years.

Because SO-FIA forest surveys are the only source of detailed data on private forest resources for extensive areas in Midsouth states, and because comparable data are gathered for public forests, they have been extremely valuable in providing information relevant to regional forest planning and management issues associated with timber production. With the passage of new laws in the mid-70's, Congress mandated that the Forest Service provide a comprehensive assessment of forest resources to include not only timber, but water, range, wildlife habitat, and recreation attributes as well.

This mandate, essentially an interdisciplinary effort, is burdened by obstacles that may not have been envisioned when the laws were passed. Interdisciplinary studies that link social and natural sciences often lack an institutional support structure; can engender incompatible priorities and perceived responsibilities among disciplines; require more time to coordinate among disciplines; have few collected works from which to gather relevant information, and even fewer journal outlets; and often lack sources for research grants and a constituency concerned with interdisciplinary issues (Heberlein 1988). Because there are few journal outlets, interdisciplinary studies are reported frequently in proceedings, rather than more widely disseminated and archived journals. My own experience leads me to believe that integrating multiple value forest surveys into an existing forest survey meets with the same types of obstacles.

I will discuss the status of multiple value integration by annotating multiple value SO-FIA research accomplishments and works in progress since the late-70's, touching on problem areas and possible solutions for the coming decade. Discussion of three subtopics follows: current conditions and trends, new approaches and additional monitoring measures, and establishment of a user constituency.

CURRENT CONDITIONS AND TRENDS

Forest survey data are in great demand by timber companies and forestry consultants. The social support structure for disseminating this information -- namely project administrators, lists of contacts, public information officers, and publications staff -- is built to provide timber-related statistics on a timely basis. Data needs have grown since the 1930's to include not only the current wood supply for a region, but also specific concerns about wood quality, availability from landowner groups, road accessibility, regeneration, and timber harvesting trends. At the National level, data needs are directed by the desire to maintain continuity in successive assessment reports, as well as to address user constituency needs for standardized statistics from different parts of the country.

In regions where the demand for timber production and other forest values conflict, private individuals and public advocacy groups have become sophisticated in articulating their demand for comprehensive planning, including multiple value forest resource information. Yet these individuals and groups frequently are unable to translate their demands to inventory data needs. Their needs are not readily answered by an ongoing, established forest survey -- a survey that was originally designed to answer questions about timber supply. Regional and National demand for other forest values is variable, as it is weakly defined in terms of priorities and statistical requirements. Organization, direction, and standardization of information needs at the National level are limited, although there has been some progress in this regard (Schlatterer and Lund 1984, Lund 1986).

Within SO-FIA's region, studies of tree biomass (e.g., Rosson and Thomas 1986) and incidence of insect and disease damage to timber resources (e.g., Mistretta and Bylin 1986) were recognized as logical and statistically definable extensions of the existing sample design. Tree biomass and tree damage analyses continue to be studied and reported at regional levels of aggregation. Coordinated efforts for National compilation are underway. Progress in the other disciplines (hydrology, soils, range science, wildlife, and recreation) is slower, but has been growing in recent years.

An important challenge for the current system of data collection, storage, and retrieval is in reorienting staff, reports, office and field procedures, forest survey data, and sampling methodology toward these other disciplines. Today's forest inventory specialist needs to be aware of the issues, methods, and literature in ecology, hydrology, range science, wildlife management, and recreation disciplines, as well as timber measurements if he/she is to be effective in implementing multiple value forest surveys. Office and field staff at the SO-FIA are being trained, or have been trained, in several of these other disciplines.

But since no one person or work unit can afford to have the expertise in all these disciplines, the SO-FIA Unit has taken a broad-brush approach to data presentation and analyses in "nontimber" (i.e., water, range, wildlife, and recreation components or values of forests) reports. The objective is to attract National, state, and university policy analysts and consultants with expertise in nontimber disciplines toward a more in-depth examination of inventory data. Graphics help reach a diverse audience not well versed in forest survey statistics. A cornerstone in our approach is an in-place, interactive data-base management system that helps answer detailed questions about issues relevant to the other disciplines.

There are gaps in our data collection effort, however. At the present time, detailed forest vegetation measurements are published and made readily available only for timberland plots -- not in designated forested wilderness areas, forested urban areas, or nonforested areas (FIA Staff 1988a). As such, habitat delineation for forest-dwelling species that utilize nontimberland areas is incomplete. Standardized bulletins and tabular information programs generally are focused on timely reporting of timber production statistics, rather than other resources. Trend information also is lacking for some of the added measures (e.g., FIA Staff 1988b), as many have been sampled only once since 1980 in Midsouth states. Data from other agencies are used to make comparisons of forest survey data with wildlife populations, soils, and recreation facility inventories, but quality, quantity, and level of detailed information vary from state to state.

Despite the above limitations, SO-FIA has succeeded in delineating habitats and trends for a few wildlife species and forest ecosystems at the state and regional level in response to issues gleaned from the ecological and forest recreation literature. At the regional level, Midsouth forest inventory data have been used to tie detailed forested area estimates with Soil Conservation Service (SCS) cropland estimates by county for multi-county physiographic units. The integrated data have been used to project land use changes (Alig et al. 1988), deer and wild turkey densities (Flather et al. 1989), and forested area, forest type, and ownership class changes (Alig et al. 1986) in association with U.S. Census county estimates of per capita income and population for Midsouth states.

An endangered forest-dwelling species, the red-cockaded woodpecker (RCW), has received attention by the press, environmental groups, and timber interests in the South, especially in east Texas. The RCW's habitat and trend estimates (Lennartz et al. 1983; Rudis 1988b: 8; Rudis 1988c: 6) have received close scrutiny in SO-FIA publications. Established criteria help delineate potential habitat and are based on detailed research from specific areas on known RCW sitings. Application of such criteria to FIA data without independent validation requires an important assumption. One must assume the generalizability of detailed research toward the larger area represented by FIA surveys. Despite this assumption, linkage of such criteria with an extensive-area data base does provide estimates about which opposing interests can argue. Because of the potential for litigation, the text is worded carefully, statistical confidence is noted, and -- most importantly -- estimates reflecting other opinions about habitat classification are provided.

Based on surveys conducted since the 1930's in the Midsouth states, we have noted declines in forested wetlands (Rudis and Birdsey 1986, McWilliams and Rosson, 1989). Subtle species and forest type shifts suggest a greater decline in moist, poorly drained forests than the wetter sites. In the Lower Mississippi Delta, historic logging for valued oaks, increases in water impoundments, and demand for soybean acreage over the years have resulted in increases in cull volume and changes in species composition toward cypress and away from overcup oak on remaining forests (Rudis and Birdsey 1986).

SO-FIA reports note dramatic declines in the fire-dependent longleaf pine forests that once dominated the southern portion of the Southern Coastal Plain in East Texas (McWilliams and Lord 1988: 2-4; Rudis 1988c: 6), Louisiana (Rudis 1988b: 18-20), and Mississippi (Kelly and Sims 1989). Our reports note that longleaf pine acreage has declined over the years and that regeneration to longleaf pine has been relatively rare. Several endangered species -- the red-cockaded woodpecker, the gopher tortoise, and several grass species -- are not exclusive to longleaf pine habitats. Nevertheless, several scientists (Seagle et al. 1987; Lohoefener 1981; and Steve Orzell, Texas Natural Heritage Program, Austin, TX, personal communication) suggest that the decline in natural stands of longleaf pine habitat explains part of the decline in the populations of these species.

Characterization of the regional landscape appears to be a strength of the existing sample design, as its use provides regional insights in studies of biogeography and landscape ecology. Individual species plot distributions, a relatively simple approach to spatial data analysis, provide thematic maps which highlight the physiography, biogeography, and human influences of selected species within the region. Maps of forested plots that highlight adjacent areas (e.g., water, urban land use, forested land cover) help one to grasp the regional context -- the landscape ecology -- of forest resources. For example, remote or relatively contiguous forest cover and roadless areas have been noted as habitat for black bears and other species in need of seclusion, as natural area buffers for designated wilderness areas, and as potential wilderness recreation areas (Rudis 1986).

NEW APPROACHES AND ADDITIONAL MONITORING MEASURES

On a more routine basis, we have incorporated water, soils, range, wildlife, and recreation components of forest resources into Midsouth states' reports and field measures when possible. New approaches and measures have been required to reorient forest survey data compilation and analyses toward issues sought by individuals trained in other disciplines and agencies responsible to other types of constituents.

One approach has been to provide detailed information on oak species and noncommercial species, and to examine mast species by diameter class, physiographic region, and by county (Rudis et al. 1984: 25; Rudis 1988b: 6; Rudis 1988c: 11). Additional information also is provided on dead trees as a way to inventory potential habitats for cavity excavators and cavity-nesting wildlife. Forest area, ownership, and forest type statistics have been aggregated by human population planning districts to help land development agencies consider forest resources when planning for projected population growth (Rudis 1988b: 17; Rudis 1988c: 19). Data also are presented or made available at the county level to maximize flexibility in assisting planners and agencies when compiling forest statistics for their districts.

New field methods to integrate water, range, wildlife, and recreation values with SO-FIA forest survey data have been developed. These include methods to obtain consistent estimates of understory vegetative cover (Popham and Baker 1987) and screening (Rudis 1985b). Understory vegetative cover tied to forest surveys has been used to make projections of forage to estimate livestock potential (Joyce and Baker 1987). Screening measures have been associated with scenic beauty (Ruddell et al. 1989, Rudis et al. 1988) and cover for deer and wild turkey. The importance of several forest survey parameters -- proximity to roads, urban areas, and water -- has been associated with recreation use and the Recreation Opportunity Spectrum (Rudis 1983, 1985a, 1987). With little additional work, a classification scheme for designating a range of recreation potential from attributes of sampled plots can be developed from the conclusions in Rudis (1987). Evidence of litter, one of several human intrusions noted in recent forest surveys, is an attribute which limits primitive recreation opportunities. Because "evidence of litter" is part of the extensive-area data collected, our survey has been able to provide facts related to pollution and the recycling of disposable containers that have become important issues in the past decade (Rudis 1988a, Rudis 1988b: 14, Rudis 1988c: 24).

Field measures added in the survey include inventories of livestock use, standing dead trees, human intrusions (logging, evidence of human uses), fire evidence, signs, fences, and neighboring land uses (proximity to urban and built-up land, agricultural land, and other forests) (FIA Staff 1988a). Issues such as potential habitats for wildlife that use snags, fire potential of the wildland-urban land interface, access, and vulnerability of forest conversion to other land uses have been addressed in the most recent nontimber report (Rudis 1988c). As states are resurveyed, such data will grow in importance. Remeasurement of added field measures will enable agencies to gain a handle on the direction and rate of change.

Our staff has pilot-tested a number of measures that address information needs related to soil productivity. The need for improvements in site productivity estimates has been recognized for a long time. In central Tennessee, the Smalley system of classification (Smalley 1980) is being tested to assess its value in addressing site productivity by comparing forest land classification estimates with those from historic records of forest survey data on growth of stands (John Rennie, personal communication, University of Tennessee, Knoxville, TN). In North Louisiana, SCS soil scientists have cooperated with SO-FIA by taking soil samples on SO-FIA plots to correlate soil series and productivity estimates. In Arkansas, land use and cover classified from satellite photography has been verified on the ground by SO-FIA field personnel to aid in independent area estimation conducted by the SCS. In this way, detailed soil and nonforest area measures (productivity, erosion potential, cropland area, etc.) conducted by the SCS one day may be integrated with the forest survey effort in a future geographic information system (GIS) and sampled-area data base.

Ordinal estimation of range resources has been based on surveys of limited areas. Early studies on the intensity of grazing in west Louisiana suggested that livestock use in southern forests was considerable -- over 50 percent of the timberland had been grazed (Sternitzke and Pearson 1974). Several surveys recorded potential browse and browse utilization in percent for west Louisiana, Tennessee, and Alabama. One study extended understory biomass estimates in south Alabama to plots where no measurements were taken, and projected understory biomass with changes in timber volume (Joyce and Baker 1987).

Nominal estimation of range utilization currently is based on the occurrence of livestock use evidence in Midsouth forests. Such an inventory provides wildlife agencies and others with an estimate of acreage, forest type, ownership, and approximate location of forested ecosystems where competition between wildlife and livestock needs can occur. Estimates of livestock use vary by region: 8 percent in Alabama (Rudis et al. 1984), 23 percent for East Texas (Rudis 1988c: 5), 38 percent for East Oklahoma (unpublished). Examination of the mapped data corroborate an independent study by Byington and others (1983) that livestock use of southern forests is a localized phenomenon -- occurring predominantly in pasture-dominated areas of states and in farmer-owned forests where the local history of forest land use includes livestock grazing of forested land.

Inventories of lesser vegetation associated with forests have been conducted in western Louisiana (Pearson and Sternitzke 1974), Tennessee, and Alabama (Joyce and Baker 1987). However, due to a variety of problems, comprehensive monitoring across all states has not been attempted. At this time, field measures are being considered to optimize cost-effective measures that limit observer variability and account for seasonal differences. Objectives are to establish field measures that estimate understory biomass and screening of vegetation, and to classify ecological communities -- i.e. seral stages and plant associations -- for range and wildlife habitat delineation. To help establish National standards, linkage of local field methods and classification schemes with National-level efforts, such as those being pursued by the Nature Conservancy, is also an important objective.

A survey of uncommon, rare, and endangered species was made in Arkansas at the request of the Arkansas Natural Heritage Program. When located on or near sampled plots, SO-FIA personnel were to identify and record potential red-cockaded woodpecker nesting sites, and record the presence or absence of 14 plant species (FIA Staff 1987). Of the 3,033 timberland plots, 90 plots (3 percent) contained uncommon, rare or endangered species. There were too few records of any one species to say anything statistically about the data collected. Because anonymity of exact locations was desired to retain continuity in the existing sample design, SO-FIA provided only general location parameters -- i.e. township, range, and section. Results of the survey yielded one previously unknown site where the endangered yellow lady's slipper might occur. Observations for 8 other species confirmed the continuing presence of these species in known areas of the State.

Integration and analysis of other inventory information from states, the National Forest System, and other federal agencies is an ongoing activity of the SO-FIA Unit. State agencies frequently provide data from their inventory efforts at little or no cost. The most versatile integration of other data sources with SO-FIA's data-base management system is at the county-level. Satellite remote sensing information has provided within-county information on land use and land cover for limited areas (Teuber 1987). Progress in these areas is encouraging and should prove fruitful for nontimber applications associated with juxtaposition of cover types, complete enumeration of forest area, and linkage with SCS soils data, U.S. Bureau of the Census population statistics, and state recreation inventories in a future GIS.

ESTABLISHMENT OF A USER CONSTITUENCY

SO-FIA's mission generally is aimed toward a consensus about what should be done. Just as with any public agency, priorities tend to shift toward issues of concern to a user constituency: knowledgeable groups and individuals able to fund special studies, and those with political influence. Hence the process by which SO-FIA determines multiple value information needs associated with forests becomes relevant.

Prior to surveying a state, the SO-FIA Unit cooperates with state forestry organizations and state-level National Forest System administrators to obtain support and encouragement for the data obtained. Field assistants, office analysts, vehicles, and funds sometimes are provided as a cooperative effort, principally to increase the speed at which surveys are conducted and reported. Instruction in log-grading, identification of trees with insect and disease damage, and orientation of field personnel to local features are provided by regional USDA-Forest Service managers and state forestry officials. Long-term support comes from federal and state legislatures, from the USDA-Forest Service research administration in Washington, DC, and from special-interest groups. Technical advisory groups are organized to advise the SO-FIA Unit in periodic reviews of its mission, priorities, and research accomplishments. These groups usually consist of individuals in public agencies, forest industries, forestry departments in land-grant universities, and forestry consultants.

Verbal and written support for multiple value surveys has come primarily from other FIA units, the National Forest System, and university researchers engaged in similar studies. Others that have recognized the importance of the multiple value forest survey effort include administrators of Alabama's TREASURED Forest program, and others sensitive to the need for a balanced overview of forest resource values. At times, individuals within technical advisory groups include those with expertise or interests in nontimber issues.

Cooperation from other agencies and disciplines, e.g. Fish and Wildlife Commissions, Planning Departments, Office of Recreation and Parks, etc. is variable. Our experience is that personnel in other agencies and disciplines are unaware of the task we perform or the data we generate. As a result, they frequently have limited knowledge of and experience with extensive area data and the potential for multiple value forest surveys. The informed social network of users outside of traditional forestry disciplines has been very narrow. Often we have found it useful to explain what we do, provide representative information related to their interests, and solicit suggestions on how we might serve their needs.

We are continuing "outreach" efforts by expanding the mailing of our reports to individuals, university research departments, non-profit organizations (such as state-level conservation groups), and editors of nature conservation and sportsmen magazines. A recently organized task force of conservation organizations has been formed to review Southern and Southeastern Forest Experiment Station research activities (Lewis 1988). Considerable time is spent in explaining our mission to solicit cooperative research projects and issues of concern that might be addressed in our reports. Contacts include individuals in universities and Natural Heritage Programs in the states we survey, as well as government agencies in planning, wildlife management, and tourism. At the regional and National level, these include: The Nature Conservancy, National Wildlife Federation, The Conservation Foundation, and The Wilderness Society.

An expanded social network also is fostered by incorporating or referencing state agency inventory efforts and conclusions where appropriate within our forest resource assessment reports. Comparison of their information with FIA data provides another view of natural resources and issues associated with forests. Our referencing of their recreation activity and supply statistics (e.g., Rudis 1988c: 22-23) has been useful to these agencies by enabling them to reach an audience of regional forest planners, forestry consultants, and forest landowners.

DISCUSSION AND CONCLUSIONS

There has been substantial progress in the development of multiple value forest surveys in the Midsouth states, particularly with the existing sample design, forest area trends, tree characteristics, and historic records of permanent plots. Greater progress in integrating water, range, wildlife, and recreation into forest surveys could be achieved if a number of obstacles are overcome. A few of these obstacles are: (1) the relatively high cost of field observations not associated with the existing sample design of widely-scattered 1-acre permanent plots, (2) the small network of experienced personnel in other agencies and disciplines that are familiar with forest survey statistics, (3) the scarcity of basic research and inventories on water, range, wildlife species, and recreation that can be linked feasibly and statistically to ongoing forest surveys of extensive areas, (4) limited trend information for much of the nontimber data available, and (5) the lack of substantive direction on data needed for National assessments of multiple use forest inventory information.

The SO-FIA effort currently focuses most of its annual budget on maintaining the resurvey of forested permanent sample plots through field measurements. However, efforts are being explored to expand the sample to accommodate inventories for other resources (specifically lesser vegetation identification, nontimberland influences, and larger sample areas needed to characterize wildlife habitat). Specific sampling designs include subsampling within existing 1-acre plots and remote sensing of larger areas with satellite imagery. In order to integrate data gathered from other disciplines, compatible geo-referenced data for all forest survey estimates will be needed. Flexibility in data analysis (e.g., by-county estimates of deer and turkey (Flather et al. 1989)), acceptance and cooperative funding of other sampling designs more suited to other resources assessment, elimination of gaps in data compilation (routine inclusion of comparable data from forested wilderness and nontimberland with trees), and standardization of natural resource terms and definitions, should provide increased opportunities for integration.

Our "outreach" efforts -- conveying the message about who we are, what we do, and how we can help land use planners and the other agencies and disciplines -- have had mixed results. Successes, measured in terms of verbal and written recognition of the importance of survey data, are derived from the "hot" issues -- scarcity of remote forests (Rudis 1986), red-cockaded woodpecker habitat (Lennartz et al. 1983), forested wetlands (Rudis and Birdsey 1986; McWilliams and Rosson, 1989), and littering (Rudis 1988a). Map displays, well-prepared graphics, and clear statements about statistical assumptions have been far more effective than presentation of data in tables. Tabular data in resource reports, without accompanying discussions of relative scarcity or trends, have yielded limited recognition.

I must conclude that focused research and reporting efforts directed toward the "hot" issues of the day in widely-distributed journals and popular magazines can establish an important source of visibility and a constituency in soils, water, range, wildlife, and recreation disciplines. Additional consideration given to map and graphic displays should help as well. Identification, analysis, and dissemination of existing FIA information in well respected journals should increase its credibility and visibility in the scientific community of other disciplines. In addition, any well publicized, cooperative funding program that fosters interdisciplinary efforts to study regional forest resource issues should further the mission of FIA Units by providing opportunities for involvement and information exchange.

National efforts to compile and analyze interdisciplinary inventory data, and to promote further data collection by specifying additional needs, have been influential in broadening the focus of forest surveys toward the other disciplines. Greater progress can be achieved if there are local, regional, and National organizations that represent a user constituency for the data collected. The first step to get this task accomplished appears to be to inform these organizations about the potential of forest surveys to address multiple use forest resource issues. The second step appears to be to involve these organizations in data compilation and analysis whenever feasible.

REFERENCES

Alig, R.J.; Knight, H.A.; Birdsey, R.A. 1986. Recent area changes in southern forest ownerships and cover types. Research Pap. SE-260. Asheville, NC: U.S. Dep. Agric. For. Serv., Southeastern Forest Experiment Station. 10 p.

Alig, R.J.; White, F.C.; Murray, B.C. 1988. Economic factors influencing land use changes in the South-Central United States. Research Pap. SE-272. Asheville, NC: U.S. Dep. Agric. For. Serv., Southeastern Forest Experiment Station. 23 p.

Byington, E.; Child, D.; Byrd, N.; Dietz, H.; Halverson, S.; Pearson, H.; Horn, F. 1983. Management of southern U.S. farms for livestock grazing and timber production on forested farmlands and associated pasture and range lands. Morrilton, AR: Winrock International. 60 p.

Flather, C.H.; Hoekstra, T.W.; Chalk, D.E.; Cost, N.D.; Rudis, V.A. 1989. Recent historical and projected regional trends of white-tailed deer and wild turkey in the southern United States. U.S. Dep. Agric. For. Serv. Gen. Tech. Rept. RM-172. Ft. Collins, CO: Rocky Mountain Forest and Range Experiment Station. 22 p.

Forest Inventory and Analysis Research Work Unit (FIA). 1987. Other forest resources work plan, Arkansas 1987. Starkville, MS: U.S. Dep. Agric. For. Serv., Southern Forest Experiment Station. 18 p. (Appendix 35 p.)

Forest Inventory and Analysis Research Work Unit (FIA). 1988a. Forest survey inventory work plan, Tennessee 1988-89. Starkville, MS: U.S. Dep. Agric. For. Serv., Southern Forest Experiment Station. 61 p. (Appendix 42 p.)

Forest Inventory and Analysis Research Work Unit (FIA). 1988b. Other forest resources work plan, Tennessee 1988. Starkville, MS: U.S. Dep. Agric. For. Serv., Southern Forest Experiment Station. 13 p.

Heberlein, T.A. 1988. Improving interdisciplinary research: integrating the social and natural sciences. Society and Natural Resources 1(1): 5-16.

Joyce, L.A.; Baker, R. 1987. Forest overstory-understory relationships in Alabama forests. Forest Ecology and Manangement 18: 49-59.

Kelly, J.F.; Sims, M. 1989. Forest resources of Mississippi. Resour. Bull. SO- 147. New Orleans, LA: U.S. Dep. Agric. For. Serv., Southern Forest Experiment Station. 63 p.

Lennartz, M.R.; Knight, H.A.; McClure, J.P.; Rudis, V.A. 1983. Status of red-cockaded woodpecker nesting habitat in the South. IN: Wood, D.A., ed. Red-cockaded Woodpecker Symposium II; 1983 January 27-29; Panama City, FL. Tallahassee, FL: Florida Game and Fresh Water Fish Commission: 13-19.

Lewis, D.K., compiler. Unpublished report to Station Directors of the Southern and Southeastern Forest Experiment Stations, November 23, 1988. Titled: 1988 Review of United States Forest Service Southern Research by Conservation Organizations. 7 p.

Lohoefener, R. 1981. Comparison of gopher tortoise (Gopherus polyphemus) habitats in young slash pine and old longleaf pine areas of southern Mississippi. Journal of Herpetology 15(2): 239-242.

Lund, H.G. 1986. A primer on integrating resource inventories. Gen. Tech. Rep. WO-49. Washington, DC: U.S. Dep. Agric. For. Serv., Timber Management, 64 p.

McWilliams, W. H.; Lord, R. G. 1988. Forest resources of East Texas. Resour. Bull. SO-136. New Orleans, LA: U.S. Dep. Agric. For. Serv., Southern Forest Experiment Station. 61 p.

McWilliams, W.H., and Rosson, J.F., Jr. 1989. Composition and vulnerability of bottomland hardwood forests of the Coastal Plain province in the South Central United States. Forest Ecology and Management.

Mistretta, P.A.; Bylin, C.V. 1986. Incidence and impact of damage to Alabama's timber, 1983. Resour. Bull. SO-112. New Orleans, LA: U.S. Dep. Agric. For. Serv., Southern Forest Experiment Station. 20 p.

Pearson, H.A.; Sternitzke, H.S. 1974. Forest-range inventory: a multiple-use survey. Journal of Range Management 27(5): 404-407.

Popham, T.W.; Baker, R.L. 1987. Ocular and densimeter estimates of understory foliar cover in forests of Alabama. Res. Note SO-334. New Orleans, LA: Southern Forest Experiment Station. 4 p.

Rosson, J.F.,Jr.; Thomas, C.E. 1986. The woody biomass resource of Alabama. Resour. Bull. SO-228. New Orleans, LA: U.S. Dep. Agric. For. Serv., Southern Forest Experiment Station. 31 p.

Ruddell, E.J.; Gramann, J.H.; Rudis, V.A.; Westphal, J.M. 1989. The psychological utility of visual penetration in forest scenic beauty models. Environment and Behavior 21(4): 393-412.

Rudis, V.A. 1983. Dispersed recreation inventory on commercial timberland. IN: Bell, J.F.; Atterbury, T., eds. Renewable resource inventories for monitoring changes and trends. 1983 August 15-19; Corvallis, OR. Corvallis, OR: Oregon State University, College of Forestry: 214-218.

Rudis, V.A. 1985a. Timber surveys: potential for dispersed recreation resource assessments. IN: L.M. Anderson, ed. Proceedings: 6th Annual Southeastern Recreation Research Conference; 1984 February 16-17; Asheville, NC. Athens, GA: Univ. Georgia, Institute for Behavioral Research: 51-58.

Rudis, V.A. 1985b. Screenometer: a device for sampling vegetative screening in forested areas. Canadian Journal of Forest Research 15: 996-999.

Rudis, V.A. 1986. Emerging patterns in the distribution of roadless forested areas in the Midsouth. IN: Kulhavy, D.L.; Conner, R.N., eds. Wilderness and natural areas in the eastern United States: a management challenge. Nacogdoches, TX: Center for Applied Studies, School of Forestry, Stephen F. Austin State University: 265-270.

Rudis, V. A. 1987. Recreational use of forested areas by Alabama residents. U.S. Dep. Agric. For. Serv. Research Paper SO-237. New Orleans, LA: Southern Forest Experiment Station. 37 p.

Rudis, V.A. 1988a. Litter in Alabama's forests. Alabama's TREASURED forests. 7(1): 15.

Rudis, V.A. 1988b. Nontimber values of Louisiana's timberland. U.S. Dep. Agric. For Serv. Resour. Bull. SO-132. New Orleans, LA: Southern Forest Experiment Station. 27 p.

Rudis, V.A. 1988c. Nontimber values of East Texas timberland. U.S. Dep. Agric. For. Serv. Resour. Bull. SO-139. New Orleans, LA: Southern Forest Experiment Station. 34 p.

Rudis, V.A.; Birdsey, R.A. 1986. Forest resource trends and current conditions in the lower Mississippi Valley. U.S. Dep. Agric. For. Serv. Resour. Bull. SO-116. New Orleans, LA: Southern Forest Experiment Station. 7 p.

Rudis, V.A.; Gramann, J. H.; Ruddell, E.J.; Westphal, J.M. 1988. Forest inventory and management-based visual preference models of southern pine stands. Forest Science 34(4): 846-863.

Rudis, V.A.; Rosson, J.F., Jr.; Kelly, J.F. 1984. Forest resources of Alabama. U.S. Dep. Agric. For. Serv. Resour. Bull SO-98. New Orleans, LA: Southern Forest Experiment Station. 55 p.

Schlatterer, E.; Lund, H.G., eds. 1984. Proceedings of the inventory integration workshop, Portland, OR, October 15-19, 1984. Washington, DC: U.S. Dep. Agric. For. Serv., Range and Timber Management Staffs. 165 p.

Seagle, S. W.; Lancia, R.A.; Adams, D.A.; Lennartz, M.R. 1987. A multivariate analysis of rangewide red-cockaded woodpecker habitat. Journal of Environmental Management 25: 45-56.

Smalley, G.W. 1980. Classification and evaluation of forest sites on the Western Highland Rim and Pennyroyal. New Orleans, LA: U.S. Dep. Agric. For. Serv., Southern Forest Experiment Station. 120 p.

Sternitzke, H.S.; Pearson, H.A. 1974. Forest-range resource statistics for southwest Louisiana parishes. Resour. Bull. SO-50. New Orleans, LA: U.S. Dep. Agric. For. Serv., Southern Forest Experiment Station. 22 p.

Teuber, K. B. 1987. Use of LANDSAT thematic mapper data for classification of forest lands in northern Louisiana. IN: Proceedings, XI Pecora Symposium. Falls Church, VA: American Society for Photogrammetry and Remote Sensing: 451-456.

Expanding Conventional Timber Inventories into a Broader, Multiresource Inventory in the Southeastern States

Noel D. Cost, Project Leader
Forest Inventory and Analysis
Southeastern Forest Experiment Station
USDA Forest Service
Asheville, NC 28802

ABSTRACT

The Forest Inventory and Analysis Unit at the Southeastern Forest Experiment Station (SE-FIA) expanded its traditional timber inventories into multiresource inventories starting with South Carolina in 1978. At permanent sample locations, field crews gather data on timber, land base, range, wildlife, understory vegetation, recreation, water, soil, botany, ecology, and biomass. In determining what variables to measure and in interpreting results, SE-FIA drew heavily upon expertise of scientists in other units in the Forest Service and in other public and private organizations. Despite considerable efforts to advertise the availability of nontimber data, those data have not been heavily used outside the Forest Service.

INTRODUCTION

Statewide timber inventories have been conducted by Forest Survey in the Southeast since 1933. In 1974, new authority granted under the Forest and Rangeland Renewable Resources Planning Act (RPA) resulted in a broader inventory responsibility. The RPA requires extensive inventory information on all renewable forest resources for planning and program development. In response to this need, SE-FIA launched a crash techniques development program to expand the ongoing timber-oriented forest survey into a more complete multiresource inventory. South Carolina was selected as a pilot program test state, and various initiatives were taken to bring the needed expertise together to deal with nontimber resources.

From the beginning, specialists in many related natural resource disciplines were invited to review proposed field procedures. Their responses contributed to a number of additions and refinements in the inventory techniques. Particular attention was given to "link" variables, which are indicative of more than one forest resource condition or attribute. Classifications and measurements made at each sample location focused on special information needs for evaluating wildlife habitat, recreation use, range suitability, water quality, erosion hazards related to forestry practices, and the use interactions associated with numerous forest conditions.

The fundamental hypothesis of multiresource inventory is that the Nation's forests and rangelands support distinct plant communities that develop in predictable sequences. The approach of SE-FIA's multiresource inventory is to quantify and describe these distinct plant communities. For each site, origin, vegetative composition, and stage of development are noted, and the range of prospective benefits is determined. The inventory merges information on the understory vegetation and nontimber attributes with conventional information on overstory trees at the same sample locations. The inventory further quantifies and describes human alterations in the natural development of these plant communities. In essence, the multiresource inventory is designed to establish the basic ecological relationships vital to periodic assessments and evaluations.

OUR APPROACH

We feel there are two approaches to designing a multiresource inventory. The first approach is functional: each discipline inventories the resource attributes relevant to its needs. Each one collects separate pieces of information. For example, the range group would collect range data, etc. Each does its own planning, purchasing of aerial photographs, plot preparation, data collection, data processing, and analysis. The second approach is an integrated one. This method requires one group to collect all the resource data for all disciplines. The primary advantage of this method is that all inventory data are collected at the same time and at the same place on the ground, with no duplication of effort, time or expense. One disadvantage is that all groups must clearly make their data needs known to the group conducting the integrated inventory. Another disadvantage of an integrated inventory is the complexity of an undertaking in which as much relevant information as possible must be gathered on timber, land base, range, wildlife, understory vegetation, recreation, water, soil, botany, ecology, biomass, and use interactions.

Nevertheless, SE-FIA decided to expand its timber-oriented inventory into multiresource inventory by the integrated method.

The first step in planning any inventory is to identify specific information needs and decide how this information can best be gathered. In the South Carolina pilot study, four general approaches to gathering additional information were identified:

1. Measurements and observations for all resources would be taken at each of our permanent ground sample locations. Thus, our land sampling scheme would be intact, but the number of observations would be greatly expanded.
2. Existing maps produced by other organizations would be used by overlaying matching coordinate points. In this manner, two or more data sources would produce a more complete description of the area sampled. Data from topographic and soils maps, for example, would be merged.
3. Special sampling schemes would be developed using some combination of remote sensing, conventional or high-altitude aerial photography, direct aerial observation, and ground sampling to gather certain types of information more efficiently. This approach might, for example, be most efficient for measuring and characterizing patterns of land use and forest edge.
4. Some information would be obtained in essentially final form from other resource agencies. Statistics on hunting and fishing, population employment and payrolls, for example, would be obtained in this manner.

In planning the South Carolina pilot study, a primary objective was to identify the information that fell into the first category--the information that FIA was required to gather. Meetings and discussions with specialists in wildlife, range, and ecology revealed that measurements of both woody and herbaceous vegetation are essential in their work. A search of the literature did not reveal a standard method for measuring the various layers of vegetation beneath tree canopies. Most of the earlier studies did not attempt to quantify the volumes in layers, and those that did tended only to deal with the forest canopy layers. In our search we were looking for a fast method that would work in a variety of forest conditions and provide information on species composition, quantity, quality, and spatial arrangement of all forest vegetation.

To assure efficient operation, SE-FIA had to make certain restrictions:

a. All observations had to be possible year-round.
b. All work had to be done by SE-FIA field crews rather than outside specialists.
c. Operations had to fit SE-FIA sampling methods.

We found certain similarities in the ways different renewable resources could be inventoried. For example, with only minor modifications our existing computer and data management systems, aerial photographs, and coding systems could address nontimber resource inventory needs. However, our crews needed training to make additional nontimber measurements and observations at our permanent sample locations. We used the talents and skills of specialists in wildlife, range, recreation, hydrology and soils. Specialists from USDA-FS Region 8, USDA-FS Southeast State and Private Forestry and universities were called in to: (1) review data being collected and make recommendations for changes, (2) field test the feasibility of collecting new data and, (3) analyze and evaluate data collected.

ESTIMATING CURRENT CONDITIONS

There were no standard guidelines to follow in the analysis of multiresource data. One approach was to group the various data elements into subsets pertinent to a particular evaluation subject. Over the years, SE-FIA's involvement in limited studies of deer browse, hydrology, and red-cockaded woodpecker habitat provided some experience with the evaluation-subject approach. Experience gained from our studies and information from elsewhere indicated that many items tallied to evaluate timber were useful for evaluating other forest benefits.

We first identified data elements having common value to all the evaluation subjects. These elements, which we called "link variables", included items such as geographic location, forest type, stand age, stand size, stand origin, site descriptors, and ownership class. Next, more specific data elements were added to deal with specific evaluation subjects. For example, for each sample plot, items for wildlife habitat screening models include:

Number of live-tree cavities
Location of live-tree cavities
Size of live-tree cavities
Color of resin flow in pine trees
Presence of prospective den trees
Degree of browsing
Type of diversity around sample location
Presence of water in the condition being sampled
Number of remnant trees
Occurrence of wildlife cover

A list of specific items was added to each evaluation subject. For analysis purposes, designating an end-product, such as a series of data files, proved helpful in organizing data for future reference. Files were used to develop frequencies, distribution rates, relationships, and correlations among the various resources and evaluation subjects.

In the system we designed, each data file contains the basic link variables plus those data elements pertinent to the particular evaluation subject. To date we have prepared six different data files for analysis and interpretation. Files we have generated include wildlife habitat, outdoor recreation, soil and hydrology, forest range and understory vegetation. More detailed information on the history and purpose of the inventories is available.

ANALYZING CURRENT CONDITIONS AND TRENDS

Assistance from inside and outside the Forest Service was obtained in the analysis and reporting of nontimber data. Information about the upcoming South Carolina nontimber data was disseminated to individuals and groups in three ways: (1) seminars at universities, (2) field demonstration plots, and (3) work meetings for all experts and specialists. The purpose of the inform and involve phase was to explore possibilities of analyzing data and seek ways to improve future inventories.

Seminars were conducted at Clemson University, Virginia Polytechnic Institute and State University (VPI & SU), University of Florida and University of Georgia. Out of these seminars, two universities were identified to conduct research through cooperative agreements.

Scientists at Clemson assessed the potential of the new FIA multiresource inventory system to supply data useful in recreation analysis. A major objective was to determine the utility of FIA information in the identification of potential nondeveloped area for outdoor recreation. Specific uses such as hiking, camping, fishing, presence and type of trails or roads, and posted restrictions of use were all noted for sample locations. Several of the more important characteristics were used to identify potential recreation areas from those plots without recreation use. Combining recreation use plots with plots that have recreation potential indicated that up to one half of the sample plots could be used for recreation. The use of these key factors to identify certain forest conditions proved useful in the identification of potential recreation areas (Saunders 1979). The research at VPI & SU centered on wildlife habitat ranking methods. Earlier work showed that plot data from broad-scale inventories could be used to rank habitat suitability for certain animals (Lentz 1974, Lennartz and McClure 1979). A review of the literature revealed that habitat criteria were available for game animals, but generally lacking for nongame birds and animals. Through the efforts at VPI & SU a generalized approach to habitat evaluation that utilizes quantitive information on both macro and micro-habitat parameters was developed for beaver, turkey, gray squirrel, raccoon, pileated woodpecker, white-tailed deer, bobwhite quail, and ruffed grouse. By using the habitat requirement data developed by VPI & SU we can screen our broad scale inventory data and rank stands as good, acceptable, or poor for certain animals.

Work meetings with other nontimber specialists resulted in other cooperative studies. Availability of dens with entrances 1.0 inch in diameter and larger for secondary cavity users was summarized by forest type, cover type, stand age classes and physiographic classes for Florida and South Carolina (McComb et al. 1986). Snag densities were summarized by forest types and stand age classes across Florida. Snags were more abundant in oak-hickory and oak-gum-cypress stands than in longleaf pine or slash pine stands. Stands over 60 years old supported more snags than did 0- to 30-year old stands (McComb et al. 1986).

As part of validation studies of a Wildlife and Fish Habitat Relationships data base we tested the assumption that habitats possess consistent characteristics. We compared vegetative and faunal characteristics in the South Carolina Piedmont at the regional and stand scales. At the regional scale, control and test samples were similar. At smaller scales, samples of vegetative and bird communities were less similar to regional norms. Bird-community models were reasonably successful at predicting community composition at the regional scale and less successful at smaller scales. These findings suggested that habitats defined by forest type and stand condition may represent a mesoscale phenonmenon intermediate between the scale of the individual stand and that of the biome. We concluded that (1) a small number of general parameters could be used to predict a larger array of quantitative structural characteristics; and (2) the assumption of consistency of characteristics of habitats was valid in this instance (Hamel et al. 1986).

In another joint study with specialists, FIA data were used to evaluate trade-offs between timber production and protection of soil and water resources and to develop general information about the resources which can be used for regional, State and county planning (Dissmeyer and Cost 1984).

SE-FIA staff members working independently have used the multiresource data in a number of other studies. The inventory, for example, provides a data base for investigating species, forest conditions, biomass, and nongame bird habitat. The frequency of occurrence of a particular species can be related to various forest conditions and species associations (Craver 1982 and McClure et al. 1979). Substantial volumes in rough and rotten trees, stumps, tops, limbs, and saplings can add an additional 30 percent to the total tree biomass (Cost and McClure 1982, Cost and Tansey 1985, Cost 1986, 1988).

A major problem Sheffield encountered in quantifying the habitats of nongame birds was that most available habitat information is qualitative in nature rather than quantitative. To rank stands as good, acceptable, or poor habitat, quantitative guidelines were established for each habitat parameter determined from the literature. Sheffield (1981) gleaned habitat descriptions for nine nongame birds from the literature and used habitat screening models to predict the extent and condition for each species.

CONSTITUENCY DEVELOPMENT

Since completing the 1978 multiresource inventory of South Carolina, SE-FIA has continually attempted to develop a nontimber users group by disseminating information in four different ways.

On the ground examination.--Representative from other FIA Work Units, States, Forest Service Region 8 and State and Private Forestry, NASA, Soil Conservation Service, and conservation groups have visited sample plots over the last 10 years. Ground demonstrations inform many people of the techniques that SE-FIA are using, generate ideas, and establish contacts with specialists who will be helpful in the future.

Seminars.--A number of seminars have been conducted to focus on various aspects of the multiresource inventories. In 1988, Ray Sheffield and I participated in the Wildlife Habitat Shortcourse at VPI & SU to inform a group of wildlife biologists about survey methods and techniques.

Slide-tape program.--A slide program titled "Planning and Conducting a Multiresource Inventory" has been effective in transferring information. This program has been shown at least 100 times to individuals or groups. Currently, both the University of Florida and the University of Georgia use the program in courses.

Publications.--Many articles dealing with the multiresource inventory are currently available from the SE-FIA work unit. Many have been cited in this paper.

Even now, SE-FIA makes a special effort to invite individuals from State and Federal natural resource agencies, universities, forestry consultant firms, and forest industries to briefing meetings for upcoming statewide inventories. At these meetings, FIA staff members explain the purpose and objectives of broad scale inventories in order to establish working relationships with other interest groups.

CONCLUSIONS

To date, most of our efforts to get conservation organizations to use SE-FIA's nontimber information have failed. A great deal of enthusiasm centered around the South Carolina pilot study and much was accomplished. Many individuals who were initially involved have now transferred or retired. Often new individuals in conservation organizations are not aware of the broad scale inventory task that SE-FIA performs across the Southeast or do not understand the kinds of data available for analysis. I also believe that many individuals in conservation organizations prefer basic research performed on limited geographical areas to an extensive approach across many different forest conditions. At SE-FIA, we believe both kinds of research are needed. There is no question that the quality of FIA nontimber data and analysis can improve as more basic research is conducted. In the meantime, I believe FIA should continue to gather multiresource data for evaluating the full range of forest benefits.

SUMMARY

In summary, I believe the SE-FIA unit has identified an important task and made good progress toward its completion. The project will continue to look for new ways to make the multiresource inventory as useful and as available as possible. We will continue to encourage specialists in all natural resource disciplines to assist and cooperate with FIA in the evaluation and dissemination of the inventory findings.

REFERENCES

Cost, N.D., 1986: Multiresource inventories: woody biomass in North Carolina. U.S. Dep. Agric. For. Serv. Research Paper SE-261. Asheville, NC: Southeastern Forest Experiment Station. 36 pp.

Cost, N.D., 1988: Multiresource inventories: woody biomass in Virginia. U.S. Dep. Agric. For. Serv. Research Paper SE-269. Asheville, NC: Southeastern Forest Experiment Station. 35 pp.

Cost, N.D. and McClure, J.P., 1982: Multiresource inventories--forest biomass in Florida. U.S. Dep. Agric. For. Serv. Research Paper SE-235. Asheville, NC: Southeastern Forest Experiment Station. 24 pp.

Cost, N.D. and Tansey, J.B., 1985: Multiresource inventories: woody biomass in Georgia. U.S. Dep. Agric. For. Serv. Research Paper SE-248. Asheville, NC: Southeastern Forest Experiment Station. 32 pp.

Craver, G.C., 1982: Multiresource inventories--a technique for determining the determining the distribution and extent of honeysuckle or commercial forest land in South Carolina. U.S. Dep. Agric. For. Serv. Research Note SE-317. Asheville, NC: Southeastern Forest Experiment Station. 11 pp.

Dissmeyer, G.E. and Cost, N.D., 1984: Multiresource inventories: watershed condition of commercial forest land in South Carolina. U.S. Dep. Agric. For. Serv. Research Paper SE-247. Asheville, NC: Southeastern Forest Experiment Station. 26 pp.

Hamel, P.B., Cost, N.D., and Sheffield, R.M., 1986: The consistent characteristics of habitats: a question of scale. IN: Verner, J.; Morrison, M.L.; Ralph, C.J., eds. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. Oct. 7-11; Madison, Wis: Univ. Wis. Press, pp. 121-128.

Lennartz, M.R. and McClure, J.P., 1979: Estimating the extent of red cockaded woodpecker habitat in the Southeast. IN: Frayer, W.E. ed. Forest resource inventories. vol. 1. 1979 July 23-36; Dep. Forest. and Wood Sci., CO. State Univ., Ft. Collins, pp. 48-62.

Lentz, R.J., 1974: Summary of WHEP "wildlife habitat evalutation program" for the Alabama river basin. U.S. Dep. Agric., For. Serv. SA, S&PF In-Service Rep., Jackson, Miss Jackson Field Office, 64 pp.

McClure, J.P., Cost, N.D., and Knight, H.A., 1979: Multiresource inventories: a new concept for forest survey. U.S. Dep. Agric. For. Serv. Research Paper SE-191. Asheville, NC: Southeastern Forest Experiment Station. 68 pp.

McComb C.W., Bonney, S.A., Sheffield, R.M., and Cost, N.D., 1986: Den tree characterics and abundance in Florida and South Carolina. The Journal of Wildlife Management, vol. 50, no. 4, pp. 584-591.

McComb, C.W., Bonney, S.A., Sheffield, R.M., and Cost, N.D., 1986: Snag resource in Florida - are they sufficient for average populations of primary cavity-nesters? Wildlife Society Bulletin, vol. 14, no. 1, pp. 41-48.

Saunders, P.R., 1979: Results of a multiresource inventory: analysis of undeveloped rural recreation sites in South Carolina. IN: Frayer, W.E. ed. Forest resource inventories, vol. 1. 1979 July 23-36; Dept. Forest and Wood Sci., CO. State Univ. Ft. Collins, pp. 63-71.

Sheffield, R.M., 1981: Multiresource inventories: techniques for evaluating nongame bird habitat. U.S. Dep. Agric. For. Serv. Research Paper SE-318. Asheville, NC: Southeastern Forest Experiment Station. 28 pp.

Developing Integrated Forest Surveys for the Rocky Mountain States

Renee O'Brien
Range Scientist
Forest Survey
Intermountain Research Station
Forest Service
U.S. Department of Agriculture
Ogden, Utah 84401

ABSTRACT

In response to legislative direction, Forest Survey for the Rocky Mountain States is responding to the need for multiresource forest land information. In addition to expanding inventories to include previously uninventoried tree species, Forest Survey now collects information that can be used to characterize forest lands for a variety of resource values, including forage, wildlife habitat, and recreation. Efforts are under way to evaluate data already collected, develop new methodology, and increase outside involvement by cooperators and potential data users.

INTRODUCTION

The Forest Survey Unit at the Intermountain Research Station conducts forest inventories in eight Rocky Mountain States on about a 10-year cycle. All ownerships are included, although most National Forest data are provided to the unit from management planning inventories conducted on individual forests.

The Resources Planning Act of 1974 (RPA) formalized a recognized need for more information about forest lands and provided for the inclusion of all forest land resources in the process. Previously uninventoried species such as pinyon-juniper, oak, mesquite, and various minor hardwoods are now included in Forest Survey inventories.

BACKGROUND

Forest Survey has always collected some site and physiographic information to describe timber supplies, but it has collected little data describing forests as a place for animals to forage and live and for humans to visit. In the past, emphasis was on sawlog inventory with little information about successional stage of the forest and the interrelationships affecting the trees, other vegetation, and the land base.

Many traditional timber inventory data can be used for assessment of nontraditional resource values. Area and volume estimates can be used for general assessments of wildlife habitat, for example. Another traditional data item included in forest inventories is overstory canopy cover. Overstory canopy cover data have many purposes, for example, predicting potential forage production, assessing wildlife habitat, and evaluating forest pest damage. To date, traditional forest inventory data are underutilized for nontraditional resource assessments.

Early efforts to include other resource values in State reports were general summaries of statistics obtained from outside sources, not from Forest Survey data. These attempts received little or no feedback. The traditional timber inventory procedure was expanded in 1981 to include additional resource assessments. Since then, much data have been collected, only a fraction of which has been analyzed.

NONTRADITIONAL DATA COLLECTION

Understory Vegetation

A method for inventorying understory vegetation on Forest Survey plots (O'Brien and Van Hooser 1983) was initiated during the 1980 Idaho inventory. This procedure is used to assess the crown canopy cover and height layer of those individual species that cover at least 5 percent of a 1/20th acre plot surrounding Forest Survey inventory points. Canopy cover is then estimated by height group for each of four life forms: trees, shrubs, forbs, and graminoids. In this way, the general structure of the vegetation on a plot is assessed, with the predominant species in each life form being identified and estimated separately. This procedure was designed to broadly apply to wildlife, range, and watershed—not for just one specific use. The procedure has gained interest from foresters and biologists throughout the West and has been used in some site-specific inventories.

Cover is estimated on Forest Survey plots instead of biomass because biomass fluctuates even more than cover due to seasonal and annual variations. Cover is a well-accepted means of understory assessment in the West and is more easily obtained than biomass. If estimates of biomass are necessary, there is evidence that a linear relationship exists between cover and biomass for some life forms (Mitchell et al., 1987).

Habitat typing is done for each Forest Survey timberland location where classification schemes are available. Habitat type is the basic unit in classifying lands based on potential (climax) natural vegetation (Daubenmire 1952, Pfister and Arno 1980). Habitat typing is a classification system that has been developed mostly on National Forest lands but that is widely used and gaining popularity in the Western United States (the technique has limited applicability in the East). Many predictive timber management models are being developed with habitat types in mind. Habitat types reflect the integration of all the conditions of a site, making them useful for prediction for management purposes. If the successional stage of a site within a certain habitat type is known, inferences about other resource values can be made (Steele 1984). The understory vegetation procedure described above gives further information about the successional stage of a site. The understory vegetation description and habitat type are at present the basis for many of the analyses being undertaken by Intermountain Stations's Forest Survey.

A preliminary classification scheme is available for pinyon-juniper woodlands in the Southwest (Moir and Carlton 1987). Work is also ongoing in the Great Basin to classify the pinyon-juniper resource into shrub associations. Both these schemes are based more on present conditions than on potential climax. If successional pathways exist in this relatively young type, they have been obscured by severe and almost universal disturbance. Present plant associations, however, may still reflect site potential.

Wildlife Habitat

The U.S. Fish and Wildlife Service has developed the Habitat Evaluation Procedures (HEP) that generally use habitat suitability index (HSI) models and are designed primarily for environmental impact analysis and planning (Schamberger and Krohn 1982). The Forest Survey data base contains many of the basic variables required for the estimation of potential suitable wildlife habitat using such models. However, many of the models are only rough quantifications of a creature's requirements, few deal with forest conditions exclusively, and most have not been validated. One attempt to use and validate an HSI model with Forest Survey data was inconclusive, and the results have not yet been published.

There are other problems involved with using Forest Survey data to estimate area of potential suitable wildlife habitat. For example, most experts in the wildlife field assess habitat on a smaller scale using intensive, site-specific procedures. They have little or no experience with extensive, sample-based inventories. Consequently, they do not immediately accept the idea of Statewide or regionwide assessments of potential suitable habitat, especially for the common species that are the focus of many of the developed models.

Estimating potential suitable habitat for some rare or endangered species (such as the pileated woodpecker [*Dryocopus pileatus*]) from Forest Survey data is one way to generate potential user interest. Because many of the rare or endangered creatures of the forested habitats of the Rocky Mountains are tied to "old-growth" in some way, this will be one area of emphasis in the Montana and Idaho inventories. Because Forest Survey in the Rocky Mountains does not usually inventory National Forest lands, data gaps exist that make regional estimates of area of potential habitat incomplete.

Additional research needs to be done by this unit to account for adjacent conditions or edge effects. This issue is critical in the types of wildlife habitat assessments that may realistically be made. Also, using indicator species or guild requirements to monitor general trends in forested ecosystems shows promise as a good alternative to using individual species requirements. A guild is a group of species that exploit the same set of environmental resources in a similar way.

Cooperation

Some positive actions have been undertaken by Intermountain Station's Forest Survey. For example, the unit currently has a cooperative agreement with Utah State University to develop a forest successional model for prediction of plant community composition and stand structure changes over 10-year intervals for two habitat types in northwestern Montana. This agreement was initiated by Forest Survey as part of a continuing attempt to expand our expertise in providing more meaningful data to a wider audience. One specific use of this model will be to predict change in extent of area of habitat over time for two species of wildlife using HSIs. The different conditions of change will be based on prescribed and natural fire and different timber management practices. Whereas Statewide estimates of potential suitable habitat for many species may not be useful, change over time to the proportions of different conditions in a State or region due to management practices may be useful. These predictions have the potential to be one valuable use of Forest Survey data for wildlife habitat assessment.

The understory vegetation description was expanded recently so that information on preferred grizzly bear food could be recorded in northwestern Montana. These data are being collected upon special request by the State of Montana and National Forest Northern Region as a result of close contacts with State and National Forest System personnel during the Montana inventory planning process. The data will be used to help verify grizzly bear management zone maps for lands outside National Forest. This was seen as an excellent opportunity to use Forest Survey data in an applied wildlife management situation, on a popular and visible issue.

One nontraditional use of timber inventory data is estimation of snag densities from standing dead tree information. The density of snags in forests in general is under scrutiny because of their importance to so many animals, their indication of the general condition of the forests, and their apparent scarcity in many western areas (Langelier and Garton 1986, Morrison et al., 1986). As a means of comparison to Forest Survey variable plot estimates of snag densities, a count of all snags on half an acre around the first point on each Forest Survey plot was taken in Montana. These data will be analyzed when the Montana inventory is completed.

Increased involvement with nontimber experts and potential users outside of Forest Survey will require a more active effort on the part of Forest Survey to solicit these interactions.

Forage and Browse

Understory vegetation and habitat type information are also the basis for assessments of forage and browse potential on Forest Survey plots. Statistics on occurrence and average canopy coverage of plant species that are preferred as forage or browse can indicate forage and browse potential. Occurrence and significant canopy coverage of species such as broom snakeweed (*Gutierrezia sarothrae*) indicate overuse of rangeland and can likewise be estimated. Reports containing information on average herbaceous or woody cover by forest type, owner, and overstory canopy cover are planned for Arizona and New Mexico.

Additional Resource Information

Descriptive, locational observations are made at each site that are useful for describing recreational and livestock use, water quality, and soil characteristics, and for further describing wildlife habitat and browse potential. A complete list of these variables is found in any Intermountain Station Forest Survey field manual.

Three of these descriptive variables are proving to be valuable in assessments of the extent of erosion in the pinyon-juniper woodlands. Because of the history of abuse in this type involving the removal of soil-holding plant cover throughout the last century, soil erosion is a serious problem. Litter depth, percentage bare ground, and evidence of erosion are all recorded at each location. Summaries of these conditions are being planned for evaluating the current situation of the soil resource under pinyon-juniper woodlands.

CONCLUSION

Obviously, more research is needed in developing multiresource procedures and providing holistic forest resource information. A review of the past 15 years shows progress and change, but data are still largely segregated functionally, even for the nontimber resources. We need to strengthen our knowledge and information about ecological relationships between interacting forest resources with corresponding production information. The gap between sample-based and spatial data should be reduced to improve the analytical utility of Forest Survey data.

Increased involvement by experts outside Forest Survey, from other agencies and the private sector, is essential if the large data bases now available are to be used to their full potential. This exposure and involvement will require a more active effort on the part of Forest Survey at the Intermountain Station.

Biodiversity, global warming, and other environmental issues are creating additional information needs for large-scale vegetation surveys, but demand for Forest Survey information Nationally and regionally has historically had economic origins. A total woodpile approach has evolved from a sawlog-oriented beginning, but forest surveys are only beginning to provide total ecosystem information.

REFERENCES

Daubenmire, P., 1952: Forest vegetation of northern Idaho and adjacent Washington, and its bearing on concepts of vegetation classification. Ecological Monographs, 22:301-330.

Langelier, L.A., and Garton, E.O., 1986: Management guidelines for increasing populations of birds that feed on western spruce budworm. Agricultural Handbook 653. U.S. Department of Agriculture, Forest Service, 19 p.

Mitchell, J.E., Bartling, P.N.S., and O'Brien, R.A., 1987: Understory cover - biomass relationships in the front range ponderosa pine zone. Research Note RM-471. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Intermountain Research Station, 5 p.

Moir, W.H., and Carlton, J.O., 1987: Classification of pinyon-juniper (P-J) sites on National Forests in the Southwest. In: Proceedings, Pinyon-juniper Conference; 1986. Reno, NV. General Technical Report INT-215. U.S. Department of Agriculture, Forest Service, Intermountain Research Station, 216-226.

Morrison, M.L., Dedon, M.F., Raphael, M.G., and Yoder-Williams, M.P., 1986: Snag requirements of cavity-nesting birds: are USDA Forest Service guidelines being met? Western Journal of Applied Forestry, 1:38-40.

O'Brien, R.A., and Van Hooser, D.D., 1983: Understory vegetation inventory: an efficient procedure. Research Paper INT-323. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, 6 p.

Pfister, R.D., and Arno, S.F., 1980: Classifying forest habitat types based on potential climax vegetation. Forest Science, 26:52-70.

Schamberger, M., and Krohn, W.B., 1982: Status of the habitat evaluation procedures. Transactions of the North American Wildlife and Natural Resources Conference, 47:154-164.

Steele, R., 1984: An approach to classifying seral vegetation within habitat types. Northwest Science, 58:29-39.

Nonindustrial Wood Resources in the Rocky Mountain States

J. David Born
Forester
Forest Survey
Intermountain Research Station
Forest Service
U.S. Department of Agriculture
Ogden, Utah 84401

ABSTRACT

Nonindustrial species types, or woodlands, are an important forest resource in the Rocky Mountain States for a variety of wood products and other resource values. Inventory methodology has been developed in recent years to measure and describe woodland types. This methodology is available in a number of publications and papers cited, and woodland resource information is included in current Forest Survey reports.

INTRODUCTION

A companion paper by O'Brien (1989) describes several multiresource activities of the Forest Survey Research Work Unit at the Intermountain Research Station. Multiresource values, including additional "nontimber" wood products, are available from vast areas of forested lands in the Western States that are not suitable for industrial wood products and therefore have not been described in previous nationwide forest land assessments.

National assessments have recently included an area estimate for a category of "other" forest land. This category includes the area dominated by commercial species but having a potential productivity rate of less than 20 cubic feet per acre per year. These areas represent only about 10 percent of the 138 million acres of forest land in the Rocky Mountain States and often occur in high elevations and on steep slopes where scenic and watershed values far exceed other considerations.

A much larger segment of the "other" forest land category is composed of types such as pinyon-juniper, oak, mesquite, and various minor hardwoods. Most species included are in the genera *Pinus, Juniperus, Quercus, Prosopis,* and *Cercocarpus.* To public land management agencies, these types are known collectively as woodlands.

These woodlands cover about 52 million acres, or 38 percent of the total forest land area in the Rocky Mountain States, with the pinyon-juniper type accounting for over 80 percent of this area. Although National Forests contain over 60 percent of the area managed for industrial wood production, or timberland, they contain only 34 percent of the low productivity and woodland area in these States. However, much of this low productivity and woodland area is under Federal jurisdiction, being administered by the U.S. Department of the Interior Agencies, the Bureau of Land Management and the Bureau of Indian Affairs.

BACKGROUND

History

The past treatment of woodlands has had a great impact in shaping present stand conditions. From the Anasazi Indian period through European exploration and settlement of the West, woodland types have been used for food, fuelwood, posts, and building materials. Nineteenth century railroads, mining, and other development decimated large areas of accessible woodland for these products, including charcoal production and railroad ties. In the rural Southwest, and in some areas throughout the range of woodland types, fuelwood continues to be the sole source of energy for cooking and heating.

During the early part of this century, woodlands were largely ignored, except for local use. However, with reduction in wildfire because of grazing and fire control, woodlands increased dramatically in both range and density. Grazing interests pressured for removal of the trees to increase livestock forage. Stands have been chained, cabled, clearcut, crushed, sprayed, and burned, and then seeded to grass in an effort to increase forage production (Buckman and Wolters 1987). However, many of these treated areas soon returned to trees.

Management

An attitude persists in land management agencies that woodlands are rangelands and should be managed as such. Conversely, "sawlog" foresters look down on woodlands and consider woodland species to be brush or cull trees. Attitudes are slow to change, but a new school of thought is emerging for integrated land management. As specialists in the various disciplines work together, they are finding that multiresource land outputs are not only desirable but often are cost effective.

The woodland types occur between the commercial species forests and grasslands or desert and are very important for wildlife habitat. Forage value can be increased, scenic values retained, and wood products harvested by careful planning and management. A continuing demand for urban fuelwood, following a peak in 1982 from the energy crunch, reduces the need for costly eradication treatments to improve wildlife habitat or increase forage.

PROCEDURES DEVELOPMENT

During the past 15 years, Forest Survey has been a leader in developing tree measurement and inventory methodology for woodlands. A user constituency for inventory information existed in the management agencies because of a public demand for fuelwood, posts, Christmas trees, and nuts, coupled with a lack of information about the resource. Management plans now require data from all lands for integrated land management. As a result, methodology to inventory woodlands required the development of area, tree volume, and product estimation techniques.

Area Estimation

Initial estimates of forest area for various classification criteria are usually obtained from aerial photographs. This methodology works for woodlands, but because of short tree heights and low stand densities the job is difficult. Remote sensing techniques using satellite data can take advantage of spectral characteristics of various tree species and computer classification for area estimation. Landsat Multispectral Scanner digital data (Born and Pearlberg 1987) and Landsat Thematic Mapper imagery (Born 1987) have been used successfully to inventory woodland types. Infrared reflectance, indirectly measuring moisture content, has been particularly useful in separating woodland species from other trees.

Tree Measurements

A modified tree definition was needed to accommodate the unusual tree form, short heights, and the products of interest. Utilization standards and specifications for Christmas trees and posts were adopted based upon criteria in use by cooperating agencies. Forest Survey field manuals include details of the procedures (U.S. Department of Agriculture 1985).

Conventional tree measurement methods required modification to estimate volume in these short, often multistemmed, bushy trees where much of the wood volume resides in the branches. The field procedures that were developed for use on the Carson National Forest in 1975 evolved into those reported by Born and Chojnacky (1985).

Conventional diameter at breast height measurement is difficult except on the best pinyon because of branches and multiple stems. A measurement at the base of the stem or stems, which is called diameter at root collar (d.r.c.), has proven to be the most practical independent variable to model individual tree volume. For multiple-stemmed trees, the diameter equivalent for total basal area at the base of the stems is used.

The procedures include number of stems and total height as additional independent variables. A visual segmentation technique is used to obtain sample tree volumes for modeling individual tree volume. This procedure is an adaptation of methods long used for volume table development by the Southeast Forest Inventory and Analysis Project (Cost 1978). Destructive sampling during the Nevada Survey in 1981-1982 verified that the visual segmentation technique, when applied carefully, can be used to obtain good tree volume estimates during regular field sample location measurements.

Most of the individual tree volume models have been developed by Chojnacky (1985a, 1988). Ratios to discount wood volume to larger diameter limits were published by Chojnacky and Wraith (1987). Chojnacky (1985b, 1987) has also published work about woodland growth, an elusive quantity because of slow rates, total tree involvement, and difficulty in making consistent field measurements.

Forest Survey has been concerned primarily with merchantable wood volumes, but work has been done by others on total biomass (Miller et al., 1981). However, there has been little interest from users for total biomass information. Concern about global climate change, acid rain, and biodiversity may well change our approach in the future.

INFORMATION OUTPUTS

Basic woodland statistics for area, cubic volume, volume growth, and dead volume are produced in traditional style statistical tables. In contrast to statistics for industrial wood, dead volume is a desirable component for fuelwood estimates, and this component ranges from 15 to 40 percent of total cubic volume. State level reports now include woodland information, and some areas warrant separate woodland reports (Conner and Green 1988). Users like color graphics made from statistics about woodlands because trends and relationships are readily apparent.

Primary sample points located on photos or other remote sensed data sources are recorded using the UTM grid system (Grubb and Eakle 1988) at 1,000-m intervals and can be plotted at a small scale and used to form a forest type and distribution map (Born and Van Hooser 1987).

TRENDS

The strong user constituency for woodland information developed almost concurrently with the legislative mandate to inventory these lands. Forest Survey experience in making initial inventories of woodland in the Rocky Mountains and recent events indicate that the demand for information will continue.

Public lands make up a high percentage of total ownership in the Rocky Mountain States, and so Federal and State land management agencies are the primary users of Forest Survey woodland information. These agencies have strongly supported woodland inventories with funding, labor, and equipment to intensify data for management areas. National Forests have added woodland procedures to forest level management planning inventories, and Forest Survey is coordinating with the Soil Conservation Service concerning methodology for woodland soil-site correlation work.

In 1985, a pinyon-juniper conference drew a large number of participants from many fields, and three regional workshops for field training were conducted in 1988 and 1989. Forest Survey personnel participated by presenting tree sampling and measurement procedures. The exchange of ideas and experiences at these workshops is obviously helping to mold a new land ethic for woodlands and to promote the need for a holistic approach to data collection.

Fuelwood demand from woodlands is relatively strong in many areas, but economic conditions remain unfavorable for long hauls of fuelwood or chips. However, inquiries continue from potential commercial wood users for fuelwood, industrial fuel, and chipped wood products. Pinyon has been used for pulp in New Mexico, and there is a steady demand for pinyon nuts and Christmas trees. Juniper fenceposts also have a small but continuing market.

Information users continue to request more results that are site or ecosystem oriented. As new vegetation classification systems have been completed, they have been added to inventory procedures. One summary of interest to managers has been a composite classification of woodlands using slope, volume per acre, site class, and crown density. More such summaries are needed to better describe woodland resources for management planning.

REFERENCES

Born, J.D., 1987: Comparing LANDSAT thematic mapper imagery with alternative data sources for woodland inventory of the Prescott National Forest. In: Proceedings, Remote Sensing for Resource Inventory, Planning, and Monitoring. Forest Service Remote Sensing Applications Conference. 1988 April 11-15; Falls Church, VA. NSTL, Miss: American Society for Photogrammetry and Remote Sensing, pp. 177-185.

Born, J.D., and Chojnacky, D.C., 1985: Woodland tree volume estimation: a visual segmentation technique. Research Paper INT-344. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station, 16 pp.

Born, J.D., and Pearlberg, C., 1987: Forest inventory and LANDSAT MSS vegetation mapping for Arizona. In: Proceedings, Symposium on Strategies for Classification and Management of Native Vegetation for Food Production in Arid Zones; 1987 October 12-16; Tucson, AZ. General Technical Report RM-150. Ft. Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, pp. 53-58.

Born, J.D., and Van Hooser, D.D., 1987: Intermountain Forest Survey remote sensing, and geographic information. In: Proceedings, GIS '87' San Francisco: Second annual international conference, exhibits and workshops, on geographic information systems. 1987; October 26-30; San Francisco, CA. American Society for Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping, vol. 2 of two volumes, pp. 416-422.

Buckman, R.E., and Wolters, G.L., 1987: Multi-resource management of pinyon-juniper woodlands. In: Proceedings, Pinyon-juniper Conference; 1986. Reno, NV. General Technical Report INT-215. U.S. Department of Agriculture, Forest Service, Intermountain Research Station, pp. 216-226.

Chojnacky, D.C., 1985a: Pinyon-juniper volume equations for the central Rocky Mountain States. Research Paper INT-339. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, 27 pp.

Chojnacky, D.C., 1985b: Pinyon-juniper site quality and volume growth equations for Nevada. Research Paper INT-372. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station, 7 pp.

Chojnacky, D.C., 1987: Modeling volume growth for Arizona's pinyon-juniper forests. In: Ek, Alan R.; Shifley, Stephen R.; Burk, Thomas E., ed. Proceedings, IUFRO conference; 1987 August 23-27; Minneapolis, MN. General Technical Report NC-120. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station, vol. 1, pp. 247-254.

Chojnacky, D.C., 1988: Juniper, pinyon, oak, and mesquite volume equations for Arizona. Research Paper INT-391. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station, 11 pp.

Chojnacky, D.C., and Wraith, C.K., 1987: Estimating singleleaf pinyon and Utah juniper volumes for several utilization standards. Research Note INT-369. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station, 6 pp.

Conner, R.C., and Green, A.W., 1988: Colorado's woodland resources on State and private land. Resource Bulletin INT-50. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station, 50 pp.

Cost, N.D., 1978: Multiresource inventories - a technique for measuring volumes in standing trees. Research Paper SE-196. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station, 18 pp.

Grubb, T.G., and Eakle, W.L., 1988: Recording wildlife locations with the universal transverse mercator (UTM) grid system. Research Note RM-483. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, 3 pp.

Miller, E.L., Meeuwig, R.O., and Budy, J.D., 1981: Biomass of singleleaf pinyon and Utah juniper. Research Paper INT-273. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, 18 pp.

O'Brien, R.A., 1989: Developing Integrated Forest Surveys for the Rocky Mountain States. In: Proceedings, State-of-the-art Methodology of Forest Inventory, 1989 July 30 - August 4; Syracuse, NY.

U.S. Department of Agriculture, Forest Service, 1985: Arizona Forest Survey field procedures. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station, Forest Survey.

Multiresource Inventories in Alaska

Bert Mead and Vernon J. LaBau
Forester and Research Forester
USDA Forest Service, Pacific Northwest Research Station
201 East Ninth Avenue Suite 303
Anchorage, Alaska 99501

ABSTRACT

The Alaska Integrated Resource Inventory System, focusing on multiresource inventory, was conceived and developed during passage of the Alaska Lands Bill and the Alaska Native Claims Settlement act to address the broad interest displayed by numerous Federal agencies and native corporations in timber and other resources. Multiresource inventories are particularly desirable in Alaska due to the high cost of access to remote areas and the limited knowledge of all resources in these remote areas. Rather than being aimed at a particular problem or project, vegetation, soils, and wildlife measurements are collected in a descriptive manner which addresses a broad range of issues.

INTRODUCTION

The first comprehensive inventory of interior Alaska was begun in 1960 and was timber oriented. Subsequent inventories centered around available photography of forested river drainages, until the early 1970's, when the Renewable Resources Planning Act (RPA) focused attention on all renewable resources. Some of Alaska FIA's earliest multiresource work were the 1967 and 1975 Yakutat inventories evaluating moose-use on timberlands (LaBau and Fox 1984). The Susitna and Porcupine river basin inventories in the late 1970's were Alaska's first multiagency cooperative attempts at broader resource inventories (Setzer 1987; USDA 1986; Yarie 1983). See figure 1.

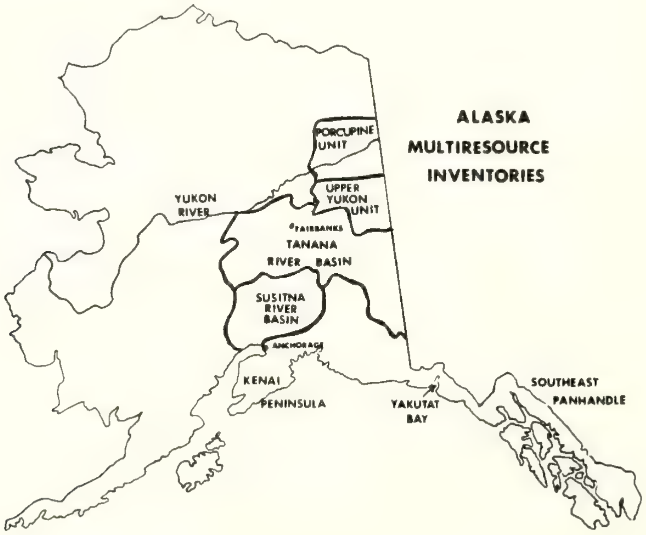


Figure 1. Map of multiresource inventories in Alaska.

DESCRIPTION OF MULTIRESOURCE INVENTORIES IN ALASKA

Since 1981, Alaska's Forest Inventory and Analysis project has been using the Alaska Integrated Resource Inventory System (AIRIS). The drivers for this system are the four-phase sampling estimators developed by the multiresource techniques project in Fort Collins, Colorado in the early 1980's (Li et al. 1984). Using multiphase sampling, we examine 20-acre (8-hectare) plots at four levels: LANDSAT multispectral scanner (MSS) data, high-altitude photography (1:60,000 scale), large-scale photography (1:3,000 - 1:7,000 scale), and ground plots. The original concept was that covariate regression estimators would permit us to develop tables of statistics by using input from all sampling levels with decreasing sampling intensity at each successive lower level (LaBau and Schreuder, 1983). Theoretically, this would permit the lowest possible sampling intensity at the expensive ground level, which is particularly appealing in remote areas of Alaska where ground plots can cost more than $2,000 each (Larson and Mead 1983). This season ground plot cost is averaging about $2,500 each.

Using LANDSAT MSS data, we classify satellite scenes by basic categories of forest, scrub, herbaceous, barren, and water. We use the data in a rather unconventional way, i.e., drawing only sample classifications of the 20-acre plots. In the process, however, huge areas are registered and classified from many different scenes (LaBau and Winterberger 1988). This process creates problems in matching scene boundaries and merging classifications from different dates, but the end-product allows us to prepare LANDSAT images of whole river basins, which are of use to resource managers and map makers.

Beginning at the second, high-altitude photo phase, we use the more detailed Alaska Vegetation Classification system developed by Viereck and Dyrness at the Institute of Northern Forestry, Fairbanks (1986). By classifying down to level II of their system, we group vegetation into cover and species-group or height classes. At the third, large-scale-photo level, vegetation is classified into forest type or nonforest species types. Using an AP-190 stereo-plotter, we obtain height and cover measurements, which are used with aerial stand volume tables to estimate stand volumes. Finally, at the ground level, we take intensive measurements on all vegetation and other plot attributes.

The AIRIS system was tested in two vastly different ecosystems, the heterogeneous 34-million-acre Tanana River basin in interior Alaska and the more homogeneous 28-million-acre southeast Alaska panhandle. In addition to employing multiphase sampling, the AIRIS system also incorporates a number of multiresource measurements. The controversy surrounding the Alaska Lands Bill focused attention on the importance of preserving or managing pristine lands by whole river basin units. Congressional legislation dictated park and refuge boundaries and State land selections so that entire ecosystems were placed under management of a single agency, with management sometimes transferred from one agency to another. This policy emphasized the importance of integrating forest and nonforest inventories as well as the unique importance of nontimber values in Alaska.

The tradeoff between timber and other resources was publicly addressed in a controversy over the proposed sale of 200,000 acres of timber in the Susitna River Basin, where FIA type maps and volume data was used by the Alaska Department of Natural Resources to delineate sale areas. The sale was temporarily withdrawn because of concerns from the hunting and fishing guide industry and the sports fishing organizations concerned about degradation of salmon spawning areas (Hulen 1987).

During the planning of the AIRIS river basin inventory system, a series of meetings were held to identify what multiresource measurements were required by our data users and to assure that our efforts were not being duplicated by other agencies. The recently published federal Fish and Wildlife Service habitat evaluation manual (1980) was reviewed to assess whether existing measurements could be used as input to the habitat models of numerous species, or whether additional data needed to be collected. In a few cases, we added additional measurements or modified existing data collection to address these still untested habitat evaluation models.

It soon became evident that there is much interest in nontree vegetation for wildlife habitat, plant ecology and forest pathology studies. One of the most versatile tools available to assess "lesser" vegetation is the horizontal-vertical (HV) vegetation profile system developed at the Southeastern Forest Experiment Station (Cost 1979). Using this system, we make extensive vegetation and habitat measurements on nonforest as well as forest plots. All plants are recorded to species level and phenological state of development is coded for each plant. When wildlife use is evident the extent of use, plant part used and type of animal using the plant are noted. Other resource measurements include length of vegetation edge, distance-to various wildlife habitat criteria (e.g. water), soil descriptions, height and cover of all vegetation, timber volumes, hiding cover, dead and downed wood descriptors, fuel loading, wildlife signs, depth to permafrost (where appropriate), and plant diversity. Preliminary discussions with Terry Daniel at a meeting on May 23, 1989 in Anchorage, indicate that some application to scenic beauty assessments is also possible with these data and the four photos taken at each HV plot (Daniel and Boster 1976).

Using variable-radius plots at 19 points evenly spaced throughout the 20-acre sample area, tree measurements are taken to estimate timberland area, volumes, growth and mortality (Grosenbaugh 1958). In addition, tree biomass estimates are made (Yarie and Mead 1982); these are particularly useful for the extensive areas of marginally productive timberland and provide a more useful estimate of this forest resource.

Understory biomass is also estimated by using cover and height values for plant layers collected on 1/40-acre (1/100-hectare) HV plots, which described vegetation composition and layering (Mead et al. 1987). Coefficients for some 120 plants and plant groups were developed under a cooperative research grant with the University of Alaska, Fairbanks, in collaboration with Dr. John Yarie (Yarie and Mead 1988a 1988b). This comprehensive information eliminated the need for labor-intensive forage and browse production plots. Because all plants, including mosses, lichens, herbs, grasses, shrubs, and trees, are identified to species whenever possible, these same plots give us descriptions of species diversity at the site.

We have also established a small herbarium and work with University of Alaska botanists to verify and map extensions of the known ranges for specific plants. The herbarium has also been used by EPA scientists and national forest personnel to make positive verification of plant specimens.

Wildlife habitat descriptions actually begin at the photo levels where distance to: water, urban development, mature trees, etc., are measured. The extent of contrasting vegetation edge on the heterogeneous plots is also measured at the photo level. The circular 20-acre large-scale photo plots are digitized, and the length of grass-forest, shrub-herbaceous, etc., edge is measured (LaBau et al. 1986). On ground plots, wildlife signs such as tracks, sightings, scat, beds, dens, etc., are also measured both in the small 1/40-acre (1/100-hectare) HV plot and throughout the 20-acre area. Hiding cover afforded wildlife is estimated across the length of the plot by using a banner of alternating black and white panels. The percent of each panel visible through the vegetation cover is estimated and recorded, thereby making the hiding cover estimates less species-specific (Griffith and Youtie 1988). A fuel-loading transect is also measured across the HV plot to determine the amount of fine and coarse fuels present (Maxwell and Ward 1981). These transects have been of interest to Tongass National Forest biologists as a way to rate marten winter habitat where dead woody material seems to play an important role. Stereophotos are also taken on the ground looking across the HV plot in cardinal directions.

Complete soil profile descriptions are taken in each type on the 20-acre plot. This is particularly important in Alaska, because only a minute fraction of the State has been mapped for soils, and this allows us to tie various soil descriptors to point or site productivity.

USES OF MULTIRESOURCE DATA

The Porcupine and Upper Yukon river basin information has been used by Doyon and Tanana Chiefs native corporations to estimate site productivity. Understory composition data was also looked at during calibration of LANDSAT spectral classes. LANDSAT classifications are being used by the U.S. Geological Survey to produce a new map product of vegetation and land cover for various areas of Alaska (Fairbanks 1989)

The Susitna river basin vegetation type maps which cover 6-1/2-million-acres at 1:63,360 scale have been used extensively by planners involved in conveyance of land into private and borough holdings via various State of Alaska programs. They used the data to estimate amount of land needed to support a subsistence lifestyle. Under one State program, most of the land remained in state ownership but small homesteads were permitted at a density determined after analysis of the understory and timber data associated with the mapped types. The maps were also used by state forestry personnel to plan state timber sales in the Susitna valley and for other land management decisions.

The HV data has also been used to develop plant association types in a area of spruce bark beetle infestation (Reynolds, 1989). This classification was in turn used as a key element for rating risk of spruce beetle attacks (Reynolds and Hard, in press). In southeast Alaska, the HV data was used to identify possible plant associations in geographic areas which had not yet been studied (Martin, 1985).

Problems and publication difficulties

We encountered some delay in generating biomass estimates from HV plot data because concurrent research to develop biomass coefficients was not yet completed. We produced height and cover displays and tables but these did not generate a lot of interest among our users, some of whom had encouraged collection of this data. The problem seemed to be that our summaries by plant groups were too general and for policy reasons, we did not produce summaries for specific ground locations. Another problem was the sheer volume of data, even at low-levels of sampling intensity. We worked with 10-20 tree species but we have a plant list of 300 to 400 species in any one river basin. This created some programming challenges when we tried to make broad-area summaries by individual plants. Yet another difficulty was the low level of sampling intensity. It was difficult to get enough plots in any classification category to make meaningful summaries and some categories had no samples, even though we sampled forest and nonforest lands at equal intensity. Summaries of these data have been made by vegetation type, forest type, and forest habitat type.

Other problems, while seemingly minor, created publication blocks. We underwent an evolutionary, informal, reporting process to achieve the right balance of detail and summarization which would be useful to our users. Another problem was that all our data and programs were originally input and output in metric units and acceptance by the general user required conversion back to English units.

We have solved many of our problems and have reached the point where we are ready to publish multiresource information which until now was only available in office reports. Two new developments have been particularly helpful. We recently received a suggestion that our detailed tables be presented on microfiche and included in the cover leaf of a publication showing a few summary tables. Second, a recent policy statement by the forest products industry encouraging publication of site specific data will allow us to publish plot level summaries identified to specific location. Such summaries have always been requested by potential users but were previously unavailable.

SUMMARY

Future Forest Inventory and Analysis efforts in Alaska will continue to be multiresource in nature, and will cover all lands within a river basin ecosystem to at least a minimal sampling intensity. Concerns about monitoring for global warming and possible changes in the vigor and health of all aspects of the ecosystem will mandate increased emphasis for comprehensive multiresource inventories.

Additionally, from an economic and sociological perspective, the conflicts are likely to increase between the demand for timber by Pacific Rim countries such as Japan, China and Korea and the demand on recreation, fisheries, and wildlife values of Alaska's land. The need for comprehensive resource assessment in Alaska and nationally will continue to grow in order to address these problems in a rational manner.

REFERENCES

Cost, Noel D. 1979. Ecological structure of forest vegetation. In: Frayer, W.E., ed. Forest resource inventories, workshop proceedings; 1979 July 23-26; Fort Collins, CO. Fort Collins, CO: Colo. State Univ., Depart. of Forest and Wood Sciences: 29-37. Vol. 1

Daniel, Terry C. and Boster, Ron S. 1976. Measuring landscape esthetics: the scenic beauty estimation method. Res. Pap. RM-167. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station. 66 p.

Griffith, Brad and Youtie, Berta A. 1988. Two devices for estimating foliage density and deer hiding cover. Wildl. Soc. Bull., vol 16, no.2, pp. 206-210.

Grosenbaugh, L. R. 1958. Point-sampling and line-sampling: Probability theory, geometric implications, synthesis. Occ. Pap. SO-160. New Orleans, LA: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station. 34 p.

Hulen, David. 1987. State moves swiftly to open Susitna Valley to timber sales. Anchorage [Alaska] Daily News. December 11.

LaBau, Vernon J. and Fox, Patricia McCartney. 1984. Use of timberlands by moose in the Yakutat area of coastal Alaska. In: Inventorying forest and other vegetation of the high latitude and high altitude regions, Proceedings of an International Symposium, SAF Regional Technical Conference, 1984. July 23-26. Fairbanks, AK: Bethesda, MD: Society of American Foresters. 249-256.

LaBau, Vernon J.; Mead, Bert R.; and Herman, David. A. 1986. Quantification of vegetation edge for the Tanana River Basin, Alaska. In: 1986 ASPRS-ACSM Fall Convention - ASPRS Technical Papers; 1986 September 28 - October 3; Anchorage, AK. Falls Church, VA: American Society for Photogrammetry and Remote Sensing. 335-341.

LaBau, Vernon J. and Schreuder, Hans T. 1983. A multiphase, multiresource inventory procedure for assessing renewable natural resources and monitoring change. In: Bell, John F. and Atterbury, Toby, ed. Renewable resource inventories for monitoring changes and trends, Proceedings of an International Conference; 1983 August 15-19; Corvallis, OR. Corvallis, OR: College of Forestry, Oregon State University: 456-459.

LaBau, Vernon J. and Winterberger, Kenneth C. 1988. Use of a four-phase sampling design in Alaska multiresource vegetation inventories. In: Satellite imageries for forest inventory and monitoring; experiences, methods, perspectives. Proceedings of IUFRO S4.02.05 meeting in Finland, Aug. 29 - Sept. 2, 1988. Hyytiala, Finland. Helsinki, Finland: University of Helsinki, Dept. of Forest Mensuration and Management. Res. Note 21: 89-102.

Larson, Frederic R. and Mead, Delbert R. 1983. Designing a comprehensive multiresource field inventory for Alaska. In: Bell, John F. and Atterbury, Toby, ed. Renewable resource inventories for monitoring changes and trends, Proceeding of an International Conference; 1983, August 15-19; Corvallis, OR. Corvallis, OR: College of Forestry, Oregon State University, Corvallis, OR. 313- 317.

Li, H.G.; Shreuder, H.T. and Bowden, D.C. 1984. Four-phase sampling estimation for the Alaska survey. In: Inventorying forest and other vegetation of the high latitude and high altitude regions, Proceedings of an International Symposium, SAF Regional Technical Conference, 1984. July 23-26. Fairbanks, AK: Bethesda, MD: Society of American Foresters. 61-67.

Martin, Jon. 1985. Key to preliminary plant associations of northern southeast Alaska: Chatham Area, Tongass National Forest. Administrative report. On file with: U.S. Department of Agriculture, Forest Service, Alaska Region, Region 10, Federal Office Building, Box 21628, Juneau, AK 99802-1628.

Maxwell, Wayne G. and Ward, Franklin R. 1981. Fuels and fire in land-management planning: Part 1. Forest-fuel Classification. Gen. Tech. Rep. PNW-131. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 12 p.

Mead, Bert; Yarie, John and Herman, David. 1987. Biomass estimates for nontimber vegetation in the Tanana river basin of interior Alaska. In: Wharton, Eric H. and Cunia, Tiberius. Estimating tree biomass regressions and their error, proceedings of the workshop on tree biomass regression functions and their contribution to the error of forest inventory estimates; 1986 May 26-30; Syracuse, New York. NE-GTR-117. Broomall, PA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station. pp. 149-153.

Reynolds, Keith. 1989. Preliminary classification of forest vegetation on the Kenai Peninsula, Alaska. Anchorage, AK: U.S. Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station. 30 p. Administrative report. On file with: Ecology and Management of the Taiga research work unit, Forestry Sciences Lab, Pacific Northwest Research Station, 201 East 9th Ave. Suite 303, Anchorage, AK 99501.

Reynolds, Keith and Hard, John. [in press]. Risk and hazard of spruce beetle attack in unmanaged stands on the Kenai peninsula, Alaska, under epidemic conditions. Forest Ecology and Management.

Setzer, Theodore S. 1987. Timber resource statistics for the Porcupine inventory unit of Alaska, 1978. Resour. Bull. PNW-RB-141. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 32 p.

U.S. Department of Agriculture, Soil Conservation Service, in cooperation with U.S.D.A. Forest Service and State of Alaska. 1986. Timber and vegetation resources of the Susitna river basin, Alaska. Anchorage, AK: U.S. Department of Agriculture, Soil Conservation Service. Unpaged.

U.S. Department of the Interior, Fish and Wildlife Service. 1980. Terrestrial habitat evaluation criteria handbook--Alaska. Anchorage, AK: U.S. Department of the Interior, Fish and Wildlife Service, Division of Ecological Services. 850 p.

Fairbanks [Topographic]. 1989. Anchorage, Alaska: U.S. Geological Survey. 1:250,000; Vegetation and land cover; colored.

Viereck, L.A.; Dyrness, C.T. and Batten, A.R. 1986. The 1986 revision of the Preliminary Classification for vegetation of Alaska. Unpublished report. Institute of Northern Forestry, Fairbanks, Alaska. 150 p. On file at INF.

Yarie, John. 1983. Forest Community classification of the Porcupine river drainage, interior Alaska, and its application to forest management. Gen. Tech. Rep. PNW-154. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station. 68 p.

Yarie, John; Mead, Bert R. 1982. Aboveground tree biomass on productive forest land in Alaska. Res. Pap. PNW-298. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Experiment Station. 16 p.

Yarie, John; Mead, Bert R. 1988a. Understory twig and foliar biomass of southeast Alaska: Biomass equations and analysis of destructive and nondestructive estimates of vertical biomass structure. Unpublished report. Forestry Sciences Lab, Anchorage, Alaska. 31 pp. On file at Forestry Sciences Lab.

Yarie, John; Mead, Bert R. 1988b. Twig and foliar biomass estimation equations for major plant species in the Tanana River Basin of interior Alaska. Res. Pap. PNW-RP-401. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 20 p.

Multiresource Inventories in the Pacific Coast States--
Progress and Future Direction

Janet L. Ohmann
Research Forester
Pacific Northwest Research Station
P.O. Box 3890
Portland, Oregon 97208

ABSTRACT

Forest Inventory and Analysis (FIA) in the Pacific Coast States began collecting additional data on nonfederal forests for multiresource assessment in the 1970s in response to changing public attitudes and new legislation. The total vegetation survey maximizes flexibility in responding to questions about habitat availability for an array of wildlife species, and for describing biological diversity at many levels of organization. New techniques for data collection, summarization, and analysis have been integrated with the existing forest inventory. FIA data are linked with habitat classification systems and models of wildlife-habitat relationships to estimate current amounts of wildlife habitats, and to evaluate habitat capability for meeting life requisites of wildlife species. Trends in habitat amounts and conditions can be monitored following remeasurement of permanent plots. Effectiveness in multiresource inventory could be improved through new and innovative strategies for reporting research and marketing products, national direction and support, and better coordination of inventories across all ownerships.

HISTORY OF MULTIRESOURCE INVENTORY ON NONFEDERAL FORESTS

Evolving Public Attitudes, New Legislative Mandates

The Forest Inventory and Analysis Work Unit (FIA) in Portland, Oregon, has been periodically assessing forest resources in Washington, Oregon, and California since 1930. Until recently the driving force behind the surveys was the need for information on timber resources--timberland area and timber volume, growth, mortality, and harvest. Since the early 1970s users of FIA data have expressed more and more interest in resource values in addition to timber. Growing public concern over environmental quality and land use allocation is reflected in legislation of the era, including the Resources Planning Act (RPA) of 1974. The Act expanded the USDA Forest Service's responsibility to assess **all** renewable forest resources--including timber, range, recreation, wilderness, water, and wildlife and fish. While philosophically FIA recognizes the need for inventory information on many nontimber resources, the Unit's primary response to the expanded mandate has been to develop approaches to inventory and monitor wildlife habitats on nonfederal forests. In this paper I discuss progress in this area, and suggest ways to improve the effectiveness of multiresource inventories.

Wildlife issues, and hence information needs for regional forest inventories, are rapidly evolving. Traditional emphasis on one or a few featured species--typically big game species--has expanded to encompass the full array of wildlife. This expanded focus has necessitated a parallel shift from censusing population numbers to monitoring availability of a diversity of habitats. The scope of issues relating to wildlife and biological diversity is also expanding, to regional and even global levels. These trends raise interesting questions and opportunities for regional forest inventories: how should biological diversity be defined, inventoried, and monitored at stand, landscape, and regional scales?

Total Vegetation Survey Maximizes Flexibility

In response to changing public attitudes and legislation FIA broadened its view of the inventory from an accounting of logs and boards to a total vegetation survey. Forest vegetation provides habitat for wildlife and is the focus of timber management activities, and therefore provides common ground for multiresource planning and management. The Unit expanded procedures for field data collection, adding a tally of nontree vegetation in 1973 and standing dead trees (snags) in 1977. In 1981 the Unit began establishing field plots in forested areas not qualifying as timberland, such as oak and juniper woodlands and chaparral (Bolsinger 1988, 1989, in press). Attributes of sampled stands are measured or observed at the most basic level that is feasible, continuous variables are used wherever practical, and stand-level attributes are calculated using computer algorithms. This approach to survey design and data collection eliminates the need to identify all information needs prior to conducting an inventory, thereby providing maximum flexibility for aggregating and summarizing data to answer a wide variety of questions. Objectivity, consistency, and repeatability are enhanced. The translation of data on vegetation composition and structure into information on wildlife habitats and elements of biological diversity is a natural extension of this approach.

New Techniques for Inventory and Analysis Needed

FIA collected the additional data perceived as pertinent to wildlife habitat assessment for several years without a clear plan for how to use it. New inventory techniques were needed before meaningful information on habitat resources could be provided; questions pertaining to sample design, field measurements, and approaches for resource analysis and data display had to be addressed. In 1981 I began the first study to develop techniques to inventory and evaluate wildlife habitat using the FIA inventory of nonfederal forests. Multiresource inventory efforts by FIA in the Pacific Coast States relate to two problem areas: quantifying and evaluating characteristics of forest resources to provide meaningful information, and developing efficient and effective technology for inventorying and evaluating the renewable resources. While work on the two problem areas is often interrelated and difficult to distinguish in practice, I will discuss each separately.

INVENTORY TECHNIQUES RESEARCH

Current Approaches Maximize Use of Existing Design and Procedures

Thus far I have developed approaches to wildlife habitat assessment that capitalize on the existing inventory design, data base, and time frame. "Piggybacking" minimizes disruption to the ongoing inventory and facilitates analysis of resource interactions, but does not meet all multiresource objectives. Keeping to the basic survey design, some data elements and procedures of data summarization have been added or modified and used in conjunction with information originally intended for timber resource analysis. The basic inventory design for FIA in the Pacific Coast States is double sampling for stratification (Cochran 1977), consisting of a primary sample of aerial photo interpretation points subsampled by a less intensive grid of field plots. Data are collected only for grid points outside parks and National Forests. Field plots consist of five sample points over a 3-hectare area. Live trees, snags, and nontree vegetation are tallied on a series of fixed-radius and variable-radius plots. Plot descriptors such as precipitation, elevation, slope and aspect, potential vegetation, and disturbance history are also recorded.

New Techniques for Data Collection and Summarization

In 1984 photo interpreters began classifying the plant community and stage of development of forested points, the data to be used to improve area estimates of wildlife habitats. While the cost of additional photo interpretation was insignificant, little gain in reliability of estimates was realized. Better results may be achieved in subsequent inventories through use of better quality and larger scale aerial photography.

Many wildlife-habitat models and systems for classifying vegetation incorporate measures of crown canopy closure. Yet there is no accurate, repeatable, and inexpensive way of measuring or estimating this attribute of stand structure as it is used in most models. For stand description for habitat analysis I adapted a procedure for estimating relative stocking that is based on individual tree contributions (MacLean 1979). Field measurements are objective and repeatable, and the approach facilitates analysis of timber-wildlife interactions by utilizing a measure commonly used in forestry applications. An informal study indicated tree stocking calculated in this manner approximates a one-to-one relationship with crown closure and is unbiased.

As FIA recently began remeasuring field plots where snags and nontree vegetation were first tallied, new questions over reconciling repeat measurements arose. Plot sizes for both snags and nontree vegetation have been held constant between inventories, simplifying field measurement and change analysis. For snag remeasurement, tally of live and dead trees is treated as a closed system: new snags are recruited when live trees tallied at the previous inventory die and remain standing; snags are lost from the tally when they shrink below minimum height or diameter thresholds, fall over, or are removed. The chief difficulty in reconciling successive measurements of nontree vegetation relates to visual estimation of canopy cover, which obscures the distinction between real change and estimator differences. Guidelines that minimize "apparent change" were developed, and recommendations on appropriate uses of FIA data on vegetation change were formed. Field methods are too inexact to quantify incremental changes that result from natural forest succession, but changes in nontree vegetation associated with major disturbances can be detected. Though the exact magnitude of change cannot be quantified, direction of change is probably accurate.

In cooperation with a university scientist I adapted line-transect procedures for sampling down logs on FIA field plots. In combination with the snag tally this provided data on characteristics and dynamics of all coarse woody debris, relevant to concerns about habitat diversity and long-term site productivity. The procedure was discontinued due to lack of funding.

Analytical Approaches to Wildlife Habitat Assessment

None of the many approaches to habitat assessment addresses all FIA objectives. An FIA approach must provide information to a wide range of users with varying interests and degrees of technical expertise. Furthermore, the resulting data base must be versatile enough to address a variety of resource issues that may arise over relatively long periods of time. As a first step in FIA efforts to evaluate wildlife habitat resources I have applied the standardized, state-level habitat classification systems and models of wildlife-habitat relationships that have been recently developed for the Pacific Coast States (Thomas 1979, Brown 1985, Mayer and Laudenslayer 1988). The analytical approach (Fig. 1) involves (1) identifying the plant community and stand condition (habitat) present on each field plot, (2) using the two-phase inventory design to estimate area of each habitat in the survey unit, and (3) linking the area estimates with wildlife-habitat models to describe potential habitat capability for wildlife species. This approach provides maximum flexibility for meeting current and future information needs since the models are standardized across broad geographic areas and the full array of wildlife species. The multi-agency support and use of these models assures a wide audience for FIA resource analyses.

Current analytical approaches to habitat assessment are aimed at estimating kinds and amounts of plant communities (species composition and structure) and special habitat elements such as snags, and evaluating how well habitat conditions meet the needs of specific wildlife species. Inventory methods and analytical approaches are discussed in Ohmann (1983, 1984, in press) and Ohmann and Mayer

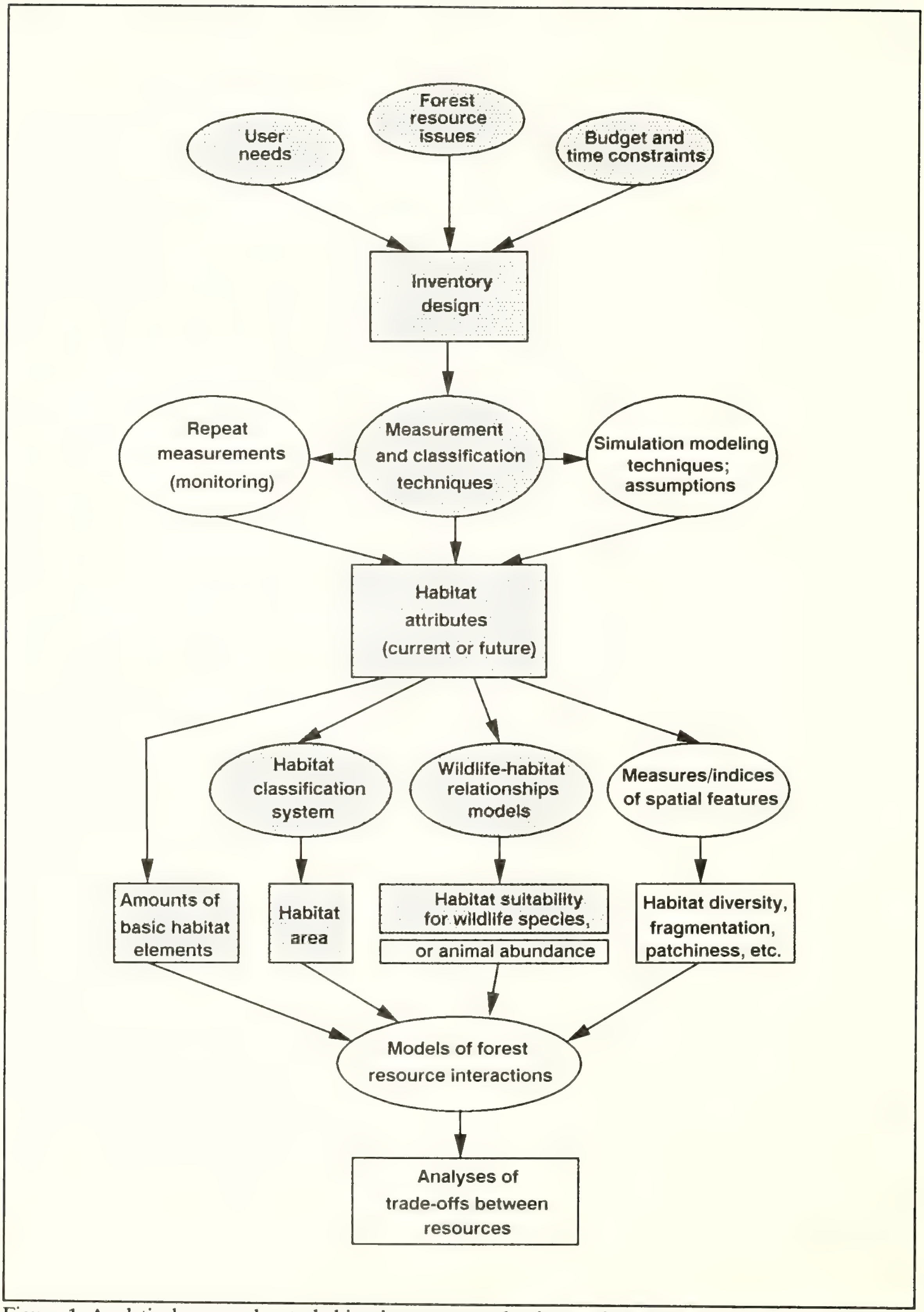

Forest resource issues
User needs
Budget and time constraints
Inventory design
Repeat measurements (monitoring)
Measurement and classification techniques
Simulation modeling techniques; assumptions
Habitat attributes (current or future)
Habitat classification system
Wildlife-habitat relationships models
Measures/indices of spatial features
Amounts of basic habitat elements
Habitat area
Habitat suitability for wildlife species,
or animal abundance
Habitat diversity, fragmentation, patchiness, etc.
Models of forest resource interactions
Analyses of trade-offs between resources

Figure 1. Analytical approaches to habitat inventory, evaluation, and monitoring (shaded portions have been applied in FIA reports).

(1987). Techniques for using the FIA inventory to monitor changes in biological diversity at several levels and to assess streamside resources have been presented at informal poster sessions (Ohmann and Bolsinger in press). Approaches to habitat assessment by FIA in the Pacific Coast States have also been presented at workshops aimed at increasing compatibility of inventory procedures among FIA units and with the National Forest System (Ohmann 1984).

Future Direction for Inventory Techniques Research

Trends such as global warming and atmospheric deposition point to the growing importance of mechanisms for defining baseline conditions and monitoring forest health and sustainability over extensive areas. Effects of such environmental changes on biological diversity at all levels of organization will receive increasing attention. With its long history of involvement and expertise in continuing forest inventory, FIA is uniquely situated to play a major role in developing approaches to monitoring biological diversity at ecosystem, landscape, and regional scales. Questions about how to estimate error, and what level of error is acceptable, will need to be addressed.

Users of FIA data in the Pacific Coast States are increasingly vocal about the need for in-place inventory information on a regional scale. Forest Survey abandoned its type-mapping work in favor of a sample-based approach in the 1950s because of the high cost of typing down to a 0.4-hectare minimum, and difficulty in quantifying change from a type-map base. New technology in remote sensing and geographic information systems offers promise for again producing in-place information about forest resources. The incorporation of information on landscape features is of particular importance in assessing habitat quality for wildlife. New techniques are needed for measuring and quantifying spatial features of forest communities, including the size, context, and arrangement of forest stands.

An important reason for developing techniques to inventory and evaluate wildlife habitats is to provide planners and managers means to evaluate multiresource implications of their decisions. Many computer models for projecting forest stands and for evaluating timber-wildlife interactions have been developed (see Marcot et al. 1988), and could be used in regional-level analyses by FIA.

ESTIMATING CURRENT CONDITIONS AND TRENDS OF HABITAT RESOURCES

Procedures for classifying plant communities and identifying stand conditions present on FIA field plots have been fully integrated into FIA operations, from inventory design and data collection to data compilation and management, including documentation of procedures. However, an analyst must monitor these steps and continue the process through data analysis and reporting of results.

New inventory techniques described in the previous section have provided information for a variety of resource reports. FIA habitat data have been incorporated in documents outlining guiding policies and programs for the state forestry departments in Oregon (Ohmann et al. 1988) and California (FRRAP 1988). The Oregon report was a joint effort of an interagency team consisting of representatives from major land-managing agencies (Bureau of Land Management and National Forest System) and the Oregon Department of Fish and Wildlife. Inventory data on current availability of plant communities, including old-growth forests, and special habitat features such as snags and logs, riparian ecosystems, winter range, and habitat diversity are presented, and implications for wildlife discussed. On a larger scale, I recently compiled an analysis of habitat conditions and trends in coniferous forests for nine western states (Ohmann 1989). The lack of compatible inventory data on wildlife habitats led me to use timber data on forest area, forest type, and stand-size included in the RPA data base. Other papers include estimates of current and projected amounts of wildlife habitats in California forests, including evaluations of habitat suitability for wildlife species (Ohmann and Mayer 1987, Ohmann in press). Estimates of timberland area near streams have been used by the State of

Oregon in evaluating proposed revisions to the Oregon Forest Practices Act. Summarized habitat data have been used by the National Forest System in compiling environmental impact statements for the spotted owl in the Pacific Coast States.

Reports of current status and trends of plant communities and stand conditions are planned for the recently completed Oregon inventory and for the Washington inventory now underway. In the future, changing amounts and conditions of vegetation and wildlife habitat resources can also be monitored using successive inventories of permanent field and photo plots, or projected into the future based on pertinent assumptions and given the analytical tools (Fig. 1).

DEVELOPING A CONSTITUENCY FOR MULTIRESOURCE INVENTORY PRODUCTS

Identifying Users of Multiresource Inventory Data

Current and potential users of multiresource inventory data have been identified through much the same means as for other inventory products. FIA personnel discuss the inventory objectives, design, and products with various groups in forums ranging from private consultations to small informal gatherings to large audiences at professional meetings. These efforts are most concentrated at the beginning of each inventory cycle, when suggestions on inventory design, data elements, and products are solicited. Results of inventories and research studies are also communicated using a variety of media, most commonly oral and poster presentations at workshops, symposia and conferences, and to meetings of professional societies, ad hoc committees, and task forces; and through publication in Station series, scientific journals, and symposium proceedings.

The most visible users of nontimber data are other government agencies--particularly state departments of forestry and fish and wildlife. Other users include research scientists and administrators (Forest Service and academia), state and national congressional delegations, the National Forest System, private forestry consultants, conservation and environmental organizations, the news media, and interested citizens. Customers are interested in both inventory techniques and resulting data and reports.

How Can FIA Be More Effective at Multiresource Inventory?

Products of multiresource inventories are beginning to find an audience in the Pacific Coast States, and levels of interest and support can be further developed in several ways. Perhaps the greatest promise, at least initially, rests with the FIA analysts and scientists themselves. The value of networking and word-of-mouth in developing local support for multiresource inventory and research cannot be overstated. FIA personnel should continue to explore new and innovative ways to package and market their products. Using a variety of outlets will assure reaching a wide range of users with varying technical backgrounds. More use of cooperative research--both extramural and interdisciplinary within the Forest Service--would also help develop support for multiresource inventories.

At the local level, FIA units could explore ways to allocate more of their limited resources to completing multiresource studies in order to demonstrate the potential value of their products. Most FIA analysts must share time for research and analysis with other primary work assignments related to FIA's inventory mandate, and also find it difficult to compete for limited technical support such as computer programming. Increased funding for multiresource inventory and research at the Station and national levels should naturally follow FIA's demonstrated ability to produce useful products and develop support of clients at the local level. Currently, top priorities of many FIA units relate to timber values, maintaining or shortening the inventory cycle, and providing quick turnaround from data collection to published reports. This has resulted in inadequate support of nontimber research.

National direction in terms of objectives, basic products, and accuracy standards for multiresource inventory and assessment are also needed. Guidelines should provide a broad framework without restricting responsiveness to local research needs. National guidance also must address coordination between inventories conducted by FIA and the National Forest System, as well as the lack of inventory data for reserved areas. In the Pacific Coast States and elsewhere, many issues of resource availability and allocation are regional in scope and not specific to a particular ownership. The current unavailability of compatible inventory data for all ownerships significantly hinders the Forest Service's ability to effectively respond to such issues, and undermines efforts to develop a constituency for multiresource inventories. Current disagreement over definitions and amounts of old-growth forest in the Pacific Northwest exemplifies this problem.

FIA is mandated to cooperate with other agencies in compiling inventories. In the Pacific Coast States FIA does not sample the major federal ownerships or areas reserved from timber harvest (primarily national and state parks). With continuing difficulty FIA obtains basic timber resource data from the National Forest System, but analyses of nontimber resources have applied only to nonfederal ownerships. In many cases relevant data for National Forests exist but are not readily accessible. The situation could be greatly improved by funding the development of systems for data management and retrieval. Also, both FIA and the National Forest System could more fully consider the other's needs in planning and designing forest inventories. However, repeated cries for such coordination have achieved limited success; changes in areas of responsibility for extensive forest inventory in the West may be needed.

The scale of FIA inventories presents another challenge in developing user support for multiresource inventories. Few potential clients are used to considering issues and problems at regional and state levels for planning and policy analysis. In addition, many users have difficulty grasping the concept of an inventory based on a sample rather than on mapped information. Furthermore, wildlife biologists, social scientists, and others are not accustomed to interpreting large amounts of tabular data that comprise the traditional FIA output. Many inventory customers have suggested making better use of graphics, maps, and prose, placing more emphasis on analysis, and translating quantitative results into qualitative information. The situation may improve as multiresource issues grow in scope and visibility, FIA efforts in multiresource analysis become better known, and user groups for multiresource data become better organized and gain political savvy.

Coordination and communication between modelers and inventory specialists also needs to improve if FIA is to benefit from increasingly powerful analytical tools being developed. For example, rarely do classification systems cover the range of plant communities and conditions encountered in FIA surveys, nor are definitions sufficiently quantitative to avoid further interpretation and second-guessing on the part of the analyst. Ideally, habitat classification systems and wildlife-habitat relationships models would incorporate habitat elements easily measured in regional forest inventories. However, very few wildlife-habitat models have been evaluated for how well they predict animal occurrence or abundance, and research in this area must continue. If easily measured elements of habitat yield poor predictions of animal occurrence, the cost and value to be gained by gathering more complex data will need to be evaluated.

CONCLUSION

At state, regional, and national levels FIA must continue to be responsive to current issues and information needs related to multiresource inventory. At the same time we must move forward in defining and building support for long-term research and monitoring of elements of wildlife habitat and biological diversity. Responsiveness to current issues wins us visibility and support; long-term research has been the traditional strength of Forest Service Research.

REFERENCES

Bolsinger, Charles L. 1988. The hardwoods of California's timberlands, woodlands, and savannas. Resour. Bull. PNW-RB-148. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station; 148 p.

Bolsinger, Charles L. 1989. Shrubs of California's chaparral, timberland, and woodland: area, ownership, and stand characteristics. Resour. Bull. PNW-RB-160. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station; 50 p.

Bolsinger, Charles L. In press. California's western juniper and pinyon-juniper woodlands: area, stand characteristics, wood volume, and fenceposts. Resour. Bull. PNW-RB-166. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Brown, E. Reade; ed. 1985. Management of wildlife and fish habitats in forests of western Oregon and Washington. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Region. 332 p. and 302 p. of appendices.

Cochran, William G. 1977. Sampling techniques, 3rd edition. New York, NY: John Wiley & Sons, Inc.; 413 p.

Forest and Rangeland Resources Assessment Program (FRRAP). 1988. California's forests and rangelands: growing conflicts over changing uses. Sacramento, CA: State of California, Department of Forestry and Fire Protection; 348 p. and appendices.

MacLean, Colin D. 1979. Relative density: the key to stocking assessment in regional analysis--a Forest Survey viewpoint. Gen. Tech. Rep. PNW-78. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station; 5 p.

Marcot, Bruce G.; McNay, R.S.; Page, Richard D. 1988. Use of microcomputers for planning and managing silviculture-habitat relationships. Gen. Tech. Rep. PNW-GTR-228. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station; 19 p.

Mayer, Kenneth E.; Laudenslayer, William F., Jr.; eds. 1988. A guide to wildlife habitats of California. Sacramento, CA: State of California. 166 p.

Ohmann, Janet L. 1983. Evaluating wildlife habitat as part of a continuing, extensive forest inventory. In: Bell, John F.; Atterbury, Toby, eds. Renewable resource inventories for monitoring changes and trends: proceedings of an international conference. 1983 August 15-19; Corvallis, OR. Corvallis, OR: Oregon State University: 623-627.

Ohmann, Janet L. 1984. Meeting wildlife habitat information needs in the Pacific Coast States--a Forest Inventory and Analysis approach. In: Lund, H. Gyde, ed. Preparing for the 21st century, proceedings of the forest land inventory workshop; 1984 March 26-30; Denver, CO. Washington, DC: U.S. Department of Agriculture, Forest Service, Division of Timber Management; p. 116-125.

Ohmann, Janet L. 1989. Status and trends of coniferous forest habitats in the western United States. In:Pendleton, Beth Giron, ed. Proceedings of the Western Raptor Management Symposium and Workshop; 1987 October; Boise, Idaho. Washington, DC: National Wildlife Federation: 38-50.

Ohmann, Janet L. In press. Wildlife habitats of California's North Coast--new techniques for extensive forest inventory. Res. Pap. PNW-RP-xxx. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Ohmann, Janet L.; Bolsinger, Charles L. In press. Monitoring biodiversity with permanent plots--landscape, stand structure, and understory species. Poster and abstract presented at symposium: Old-growth Douglas-fir forests: wildlife communities and habitat relationships; 1989 March 29-31; Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Ohmann, Janet L.; Carleson, Dan; Lee, Philip L.; Oakley, Art L. 1988. Status of forest-related wildlife and fish resources in Oregon. In: Lettman, Gary J., tech. ed.; Stere, David H., prog. dir. Assessment of Oregon's forests: a collection of papers published by the Oregon State Department of Forestry. Oregon State Department of Forestry; Salem, OR: 17-32.

Ohmann, Janet L.; Mayer, Kenneth E. 1987. Wildlife habitats of California's hardwood forests--linking extensive inventory data with habitat models. In: Plumb, T.R.; Pillsbury, N.H., tech. coords. Proceedings of the symposium on multiple-use management of California's hardwood resources; 1986 November 12-14; San Luis Obispo, California. Gen. Tech. Rep. PSW-100. Berkeley, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Forest and Range Experiment Station: 174-182.

Thomas, Jack Ward, ed. 1979. Wildlife habitats in managed forests: the Blue Mountains of Oregon and Washington. Ag. Handb. 553. Washington, DC: U.S. Department of Agriculture, Forest Service; 512 p.

State-of-the-Art Methodology of Forest Inventory: Wildlife Habitat Assessment in the Northeastern United States

Robert T. Brooks
Research Wildlife Biologist
USDA Forest Service, Northeastern Forest Experiment Station
Amherst, MA 01003

ABSTRACT

Forest wildlife habitat assessments are mandated of the U. S. Department of Agriculture Forest Service by federal legislation. This responsibility has been delegated to the regional Forest Inventory and Analysis research work units. In the northeast, habitat assessment emphasizes both within-stand and landscape habitat features. Methodologies for habitat assessment differ among units, making compilation of standard national statistics difficult. Consensus on a standard habitat survey is difficult because of the absence of federal direction and the lack of an effective constituency.

INTRODUCTION

Background

A period of strong and widespread public support for environmental issues in the late 1960's and early 1970's led to the passage of many legislative initiatives for natural resources conservation. Of specific concern to forest inventory was the Forest and Rangeland Renewable Resources Planning Act (RPA) of 1974 (P. L. 93-378). Two of the seven enumerated findings of this legislation concluded that the public interest would be served by having the USDA Forest Service, in cooperation with other public agencies, assess the Nation's renewable resources.

The comprehensive assessment was to consider "present and anticipated uses, demand for, and supply of renewable resources from the Nation's public and private forests and rangelands ...". Section 3 of the Act expanded on these findings, requiring a decadal "Assessment" including "an inventory based on information developed by the Forest Service and other Federal agencies, of present and potential renewable resources, ...". The term "renewable resources" was defined to involve those matters within the scope of responsibilities and authorities of the USDA Forest Service. In practice, renewable resources has been applied to include timber, range, recreation and wilderness, water, and wildlife and fish.

The first Assessment was completed in 1975, the second in 1979, and the third is to be completed by December, 1989. The 1975 Assessment (USDA Forest Service 1975:334) concluded:

> "Effective management and preservation of terrestrial wildlife species is dependent on management of vegetation. Much information is currently available on the vegetation of forest and range lands. However, there is a need to collect in a systematic fashion data directly useful in wildlife management such as the production of browse, forage, and mast; understory vegetation; indicator species; and foliage diversification. Special attention should be given to collecting data on habitats of threatened and endangered species and on vegetation critical to winter survival."

In effect, the Assessment concluded that there was insufficient factual data to conduct an assessment of wildlife habitat resources, basic to any decision concerning the management or use of these resources. Five years later, the 1979 Assessment developed a national data base for the wildlife assessment

(Hoekstra et al. 1979) but arrived at the same conclusion as the first Assessment (USDA Forest Service 1979:29):

> "A major constraint on understanding wildlife conditions is the lack of inventory data defining the extent of existing habitats. There is no comprehensive and quantitative inventory of available habitats for any substantial part of the Nation. Measurements of animal habitats that can be related to information on other resources are needed."

The most recent Assessment of wildlife resources made a first effort at quantifying habitat resources on a national scale (Flather and Hoekstra 1989). Since "No standard national or regional inventory exists that permits a consistent summarization of the wildlife ... resource", the habitat Assessment was made by interpretation of data collected for other purposes. While the authors theoretically defined habitat as "the availability and appropriate mix of food, cover, and water", in application, habitat was quantified by the description of the land base. Forested habitats were described by timberland forest type and stand size as respective surrogates for forest ecosystem type and successional stage.

A national, standardized forest wildlife-habitat assessment, would most appropriately come from the USDA Forest Service's Forest Inventory and Analysis (FIA) units (USDA Forest Service 1987). The seven regional research work units that constitute this program are responsible for translating legislated inventory needs into survey activities and generating survey-based forest resource information. Because of differing regional and local information needs, survey traditions, and the autonomy allowed each unit, survey techniques and data elements for habitat assessment vary considerably between units (USDA Forest Service 1987). This presents additional problems for compliation of standardized national habitat statistics. In this paper, the Northeastern Forest Inventory and Analysis efforts in forest wildlife habitat assessment are reviewed.

History

In 1977, FIA in the Northeast employed a wildlife biologist to initiate research into habitat assessment techniques (Barnes 1979, Barnes and Barnard 1979). The guidelines were broad, essentially that measurements must be easily incorporated into existing forest sampling methodologies. The mandate was to assess wildlife habitat, not faunal populations which would have been difficult in any regard, given the sampling requirements necessary to survey animal numbers. The general approach was to survey structural vegetative features, special features such as browse and mast potential and cavity trees, and landscape features. Initially these data were recorded separately from the traditional timber data set. This arrangement created difficulties in combining data for analysis and by the 1980 Maine resurvey, all data were recorded on a single combined tally sheet.

At the occasion of the Maine survey, the combined forest survey was expanded to collect data on forest soils and forest recreation, in addition to the previously mentioned timber and habitat resources. These resource data have since been dropped or simplified in the interests of time and costs. The active interest and support of a subject-matter specialist (that is, soil scientist, recreationist), are essential if the collected data are to be analyzed and reported adequately. Without such an assignment, it is difficult to identify who can or would support the collection of these data. It is critical that data recommended for inclusion in the FIA survey be supported from the beginning by a scientist or analyst with the ability, interest, and time to fully utilize and report the resource information. The relative success of wildlife habitat data collection versus that for soils and recreation is due primarily to the employment of a staff wildlife biologist.

Since wildlife habitat is an integration of many features of forest environments, both within and between stands, habitat assessment must use modeling to simultaneously access these multiple data elements (Hawkes et al. 1983). The use of habitat models will also be necessary to generate estimates of habitat units for validation studies of survey techniques. At this time, modeling and validation remain to be accomplished. Supplies of selected structural and special habitat features have been reported in Resource Bulletins (for example, Brooks 1986, Brooks et al. 1986).

ESTIMATING CURRENT CONDITIONS AND TRENDS

The Northeast FIA has recognized wildlife and wildlife habitat as an important forest resource for some time. Previously, reports on these resources used secondary information sources, generally from state wildlife agencies (for example, Powell and Considine 1982). The reports were generally limited to discussions of the popularity of wildlife, the recreational opportunities offered by wildlife resources, and the general dependency of wildlife on forest land -- similar in subject to the first national RPA Assessment of wildlife, only on a reduced geographic scale.

The premise of the Northeast FIA habitat survey was to measure a select number of forest features that had direct or interpretable relationships to wildlife habitat quality and quantity. Rather than conducting a species' specific habitat assessment, a broad range of habitat features were included to assess habitat for a diversity of species. The measured habitat features can be classified broadly as structural-, special-, and spatial-habitat features. They are measured on a variety of scales, from the species (tree or shrub), to the point, the plot, and from a fixed area surrounding the plot (that is, landscape scale) using aerial photography (Brooks and Sykes 1984). The data are sample-based and, as such, are incompatible with geographic or map-based analysis (use with Geographic Information Systems).

This information is used to quantify potential browse and mast (both soft and hard) resources, potential and utilized cavity trees including standing dead trees (snags), and landscape patterns (for example, Brooks et al. 1986). Land use and timberland area information by forest type and stand-size from the timber survey are also used. Individually, many of these data have direct value for habitat assessment, such as the numbers of trees with cavities, the trends in mast-producing tree species, and the utilization of lesser woody-stemmed vegetation as browse. The characteristics of the stand, especially stand-size and stand area and forest ownership, have important implications for habitat condition (Brooks 1986, Brooks and Birch 1986, 1988).

One can also construct models using subsets of these data for species-specific habitat assessment. This can be done by considering critical habitat elements (for white-tailed deer (Odocoileus virgineanus) Brooks et al. 1986). Using quantitative models, one can screen the survey data to quantify the potential habitat quality of each plot for individual species such as gray squirrels (Sciurus carolinensis) Barnes 1979) or cavity nesting birds (Barnes and Barnard 1979). Similar efforts have been made by the Southeastern FIA (Sheffield 1981). From these models, one can estimate acres of habitat by potential quality classes.

An important limitation of this approach, however, is that habitat classification without proof of its validity represents, at best, a consensus of expert opinion, at worst, one point-of-view. When interpretations of the data are made beyond its direct value (trees/acre), there is the implied requirement that the results be validated by independent measures of the resource being modeled. At this time, validation studies in the Northeast have been done on a limited scale and preliminary basis.

HABITAT ASSESSMENT METHODOLOGIES

Because FIA is principally a production effort with a research mandate, forest habitat survey techniques must be easily integrated into the general FIA procedure (Kaiser 1987). This includes both the sampling scheme and measurement techniques. Satisfying this requirement provides the additional benefit that the unique habitat measurements are less likely to be terminated when they are an integral and routine part of the overall survey effort.

When possible, habitat measurements are made using surrogates from the timber survey (tree mast potential from species, diameter, and crown measurements). As a second choice, relationships are developed between timber measurements and habitat features (overstory-understory relationships to estimate understory herbaceous foliage resources). When no surrogate is available and the habitat resource is poorly related to existing timber measures, recommended habitat sampling techniques must be evaluated both for their effectiveness and for their efficiency of application. Ideal techniques may have to be compromised and modified to comply with time and sample design restrictions.

CONSTITUENCY DEVELOPMENT

As detailed in the Introduction, the assessment of wildlife habitat resources is mandated by federal legislation. The legislation was not a capricious act but reflective of the mood and will of the people. This appreciation and concern for forests and forest resources in addition to timber has not abated since 1974. The opportunity exists for the Forest Service to build on this expressed public interest by supporting and encouraging a national integrated resource survey (Lund 1986).

There is little organized support for diverting fiscal resources from timber to other forest resource surveys. Since wildlife is a noncommodity resource, there is no industry-funded lobbying effort for habitat surveys. Nonprofit wildlife organizations have yet to support FIA habitat surveys. This is due in part to the demands on limited resources by more pressing issues (endangered species) but also by the failure to bring the accomplishments of the USDA Forest Service to the attention of these organizations.

Although individual states use FIA habitat survey data for wildlife planning (Maine Department of Inland Fisheries and Wildlife 1986), this use is not universal. Wildlife biologists recognize habitat as the foundation of wildlife populations, but have little experience in considering habitat resources at the spatial scale of counties, states, and regions. As wildlife planning is more rigorously mandated by federal funding agencies, demands for FIA habitat survey data will undoubtedly grow. Simultaneously, FIA wildlife specialists should participate in the Organization of Wildlife Planners (P. O. Box 7921, Madison, WI 53707), a group with interests compatible with FIA activities.

Finally, the Northeastern FIA forest ownership surveys show that forest owners consistently identify recreational use and esthetic enjoyment as primary interests in owning forest land. Wildife is an integral component of both these interests. It would be useful for FIA wildlife specialists to prepare their reports in terms meaningful to private forest owners and to produce reports that will be accessible to these owners (that is, state conservation magazines).

SUMMARY

Forest Inventory and Analysis has a long history of collecting and providing information for regional and national forest planning activities. Stimulated by federal legislation, FIA broadened its scope to consider forest resources other than wood fiber and its products. In the Northeast, FIA took a progressive course and employed a research wildlife biologist to develop, implement, and report on a forest wildlife habitat survey program. The survey considers stand- and landscape-scale habitat features. Reports have been made, and continue to be made, on the status of selected habitat resources. Preliminary efforts have been made at species-specific habitat assessments. Validation of the survey's capability to accurately assess wildlife habitat condition remains to be fully accomplished.

REFERENCES

Barnes,R.B. 1979: Wildlife habitat from a forest resource inventory: is it possible? Trans., Northeast Sec., The Wildl. Soc., vol. 36, pp. 151-159.

Barnes,R.B. and Barnard,J.E. 1979: A wildlife habitat survey as part of a multi-resource evaluation procedure. pgs. 77-89. In Frayer,W.E. Forest Resource Inventories, vol. 1. Dep. Forest. and Wood Sci., Colo. State Univ., Ft. Collins.

Brooks,R.T. 1986: Forest land wildlife habitat resources of south-central Ohio. USDA, For. Serv. Resour. Bull. NE-94. 32 pp.

Brooks,R.T. and Birch,T.W. 1986: Opportunities and constraints for wildlife habitat management on private forests of the northeast. N. Jour. of Appl. For. vol. 3, pp. 109-113.

Brooks,R.T. and Birch,T.W. 1988: Changes in New England forests and forest owners: implications for wildlife habitat resources and management. Trans. N. Amer. Wildl. and Natur. Resour. Conf. vol. 53, pp. 78-87.

Brooks,R.T. and Sykes,K.J. 1984: Sampling land use edge from aerial photographs--line transect vs. circular pattern. USDA For. Serv. Res. Note. NE-321. 6 pp.

Brooks,R.T., Frieswyk,T.S., and Ritter,A. 1986: Forest wildlife habitat statistics for Maine 1982. USDA, For. Serv. Resour. Bull. NE-96. 146 pp.

Flather,C.H. and Hoekstra,T.W. 1989: An analysis of the wildlife and fish situation in the United States: 1989-2040. USDA For. Serv. Gen. Tech. Rep. RM-178. 146 pp.

Hawkes,C.L., Chalk,D.E., Hoekstra,T.W., and Flather,C.H. 1983: Prediction of wildlife and fish resources for national Assessments and Appraisals. USDA, For. Serv. Gen. Tech. Rep. RM-100. 21 pp.

Hoekstra,T.W., Schweitzer,D.L., Cushwa,C.T., Anderson,S.H., and Barnes,R.B. 1979: Preliminary evaluation of a national wildlife and fish data base. Trans. N. Amer. Wildl. and Nat. Resour. Conf. vol. 44, pp. 380-391.

Lund,H.G. 1986: A primer on integrating resource inventories. USDA For. Serv. Gen. Tech. Rep. WO-49. 64 pp.

Kaiser,H.F. 1987: The importance of multiple use value in designing inventories. pgs. 168-171. In Lund,H.G., Caballero-Deloya,M., and Villarreal- Canton,R. (eds): Land and resource evaluation for national planning in the tropics. USDA For. Serv. Gen. Tech. Rep. WO-39.

Maine Department of Inland Fisheries and Wildlife. 1986: Planning for Maine's inland fish and wildlife, vol. I Wildlife, Part 1.1 Species assessments and strategic plans--introduction and appendices. Augusta.

Powell,D.S. and Considine,T.J. 1982: An analysis of Pennsylvania's forest resources. USDA For. Serv. Resour. Bull. NE-69. 97 pp.

Sheffield,R.M. 1981: Multiresource inventories: techniques for evaluating nongame bird habitat. USDA For. Serv. Res. Pap. SE-218. 28 pp.

USDA Forest Service. 1975: RPA the nation's renewable resources--an Assessment, 1975. U.S. Dep. of Agric. Washington, D.C. 345 pp.

USDA Forest Service. 1979: A report to Congress on the nation's renewable resources--RPA Assessment and alternative program directions, review draft. U.S. Dep. of Agric. Washington, D. C. 209 pp.

USDA Forest Service. 1987: Forest Service resource inventory: an overview. U.S. Dep. of Agric., Forest Inventory and Economics Research. Washington, D.C. 29 pp.

Accessing USDA Forest Service Forest Inventory and Analysis (FIA) Data
Moderator: Mark Hansen

Mark H. Hansen, The Eastwide Forest Inventory Database

Raymond M. Sheffield, Access to Forest Inventory Data for the Southeast

William H. McWilliams, Interactive Access to Forest Inventory Data for the South Central United States

Jerold T. Hahn and Mark H. Hansen, North Central Forest Inventory Data on a Microcomputer

Surveying With Partial Replacement
Moderator: T. Cunia

Tiberius Cunia, Successive Forest Inventories: Basic SPR Approach for Two Measurements

Tiberius Cunia, Successive Forest Inventories: Extensions of the Basic Approach

Tiberius Cunia, Successive Forest Inventories: Error of Biomass Regressions Part I: CFI With Permanent Plots Only

Tiberius Cunia, Successive Forest Inventories: Error of Biomass Regressions Part II: CFI With Partial Replacement of Plots

The Eastwide Forest Inventory Database

Mark H. Hansen
Research Forester
USDA, Forest Service
North Central Forest Experiment Station
St. Paul, Minnesota 55108

ABSTRACT

In addition to published reports, USDA Forest Service can also provide plot and tree data collected during statewide inventories in the eastern United States to those interested in further analysis. This paper describes a recently developed standard format in which such data is available. This standard format, referred to as the Eastwide Database (EWDB) structure, has been developed to provide users with consistent data on the forest resource of the eastern United States.

EWDB files can be obtained for any eastern state inventory conducted after 1988. Files for many inventories conducted before 1988 may also be available in this format. A users manual describing the EWDB is currently being prepared. This manual will provide users with the information needed to process these files. The EWDB was designed to be independent of software and hardware constraints and is structured on the relational data model supported by many database management systems.

FOREST INVENTORY AND ANALYSIS (FIA)

Forest Inventory and Analysis (FIA) is a continuing endeavor carried out by USDA, Forest Service. Its objective is to periodically determine the extent, condition, and volume of timber, growth, and depletions of the Nation's forest land. This up-to-date information is essential to frame realistic forest policies and programs. USDA Forest Service regional experiment stations conduct inventories and publish summary reports for individual states. In the East, FIA inventories are conducted by four separate projects. The interval between inventories in each state varies from five to fifteen years. The project headquarters and the states they inventory are shown in figure 1. This paper describes recent efforts by these four projects to provide data to outside users in a consistent form.

Information contained in FIA reports is developed primarily from sample plots located throughout the region. Although plot designs may vary by state, they are all designed to sample one-acre. Various arrangements of fixed and variable radius sample points are used to select sample trees on these plots. Sampling intensity also may vary by state, with one plot representing 1000 to 6000 acres. Tree measurements and other observations taken on these plots are used to describe stand conditions on the acre sampled. Estimates of conditions for a large area, such as a state, are made based on the expansion of the estimates observed on all the sample plots taken in the area using statistically sound methods.

The data obtained in FIA inventories is designed to meet a variety of needs. Published reports provide information primarily at the state level, with limited information reported for individual counties or groups of counties. Public and private forest managers use the information in these reports to help plan for future management needs and to assess the effects of past management practices on current forest conditions. Industrial planners use the information to assist in planning the direction of wood using industries. Public officials rely on FIA data to formulate public policy.

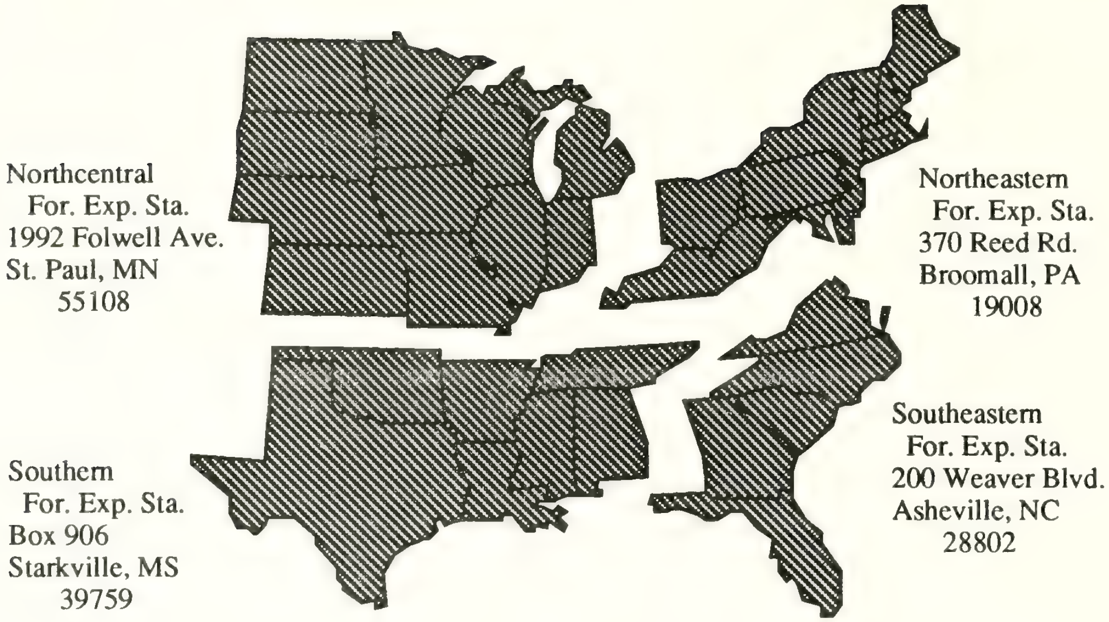


Figure 1. Eastern FIA regions and project headquarters

THE NEED FOR CONSISTENCY

Most users of FIA information find the published reports adequate for their information needs, however, others require specific information that is not contained in the published reports. These special information needs typically include:

1) Information for a geographic area other than that published.
2) Information that is consistent across States and Regions
3) Projections of timber resources based on various assumptions and models
4) Information about a specific condition and/or species

In the past, requests such as these have been handled differently by each of the four eastern FIA projects, depending on the data processing procedures and capabilities of the project. With increasing availability of electronic data processing capabilities has come an increased demand by those interested in FIA data for data files to conduct their own analysis. It has been particularly difficult to analyze data files from more than one FIA project because of inconsistencies in data processing methods between FIA projects.

One effort to improve FIA data compatible between states was the development of a set of 25 core tables that are being published in every eastern FIA inventory report. These core tables (Table 1) present basic resource information in a consistent format eastwide. The introduction of the core tables in 1985 has made it easier to compare the forest resources of areas that cross state and/or region boundaries.

A more recent effort to provide consistent FIA data throughout the East, has been the creation of the Eastwide Database (EWDB). The objective of the EWDB is for each project to produce a file for the most recent inventory of each state in its region. All files are in the same format and contain all of the data items needed to produce the core tables, plus other common data items collected by all four projects. Items which are not available from all four projects are not included in the EWDB. The EWDB data files are produced for each State as part of the routine processing of the inventory data and are available to any interested party at a reasonable cost ($500 per State).

Table 1. Eastern FIA core tables

No.	Table title
1	Area of land by county & major land use class
2	Area of timberland by county & owner class
3	Area of timberland by county & forest type
4	Area of timberland by county & stand-size class
5	Area of timberland by county & site class
6	Area of timberland by county & stocking class of growing-stock
7	Area of timberland by forest type & owner class
8	Area of timberland by ownership & stocking class of growing-stock trees
9	Area of timberland by forest type & stand-size class
10	No. of all live trees on timberland by species group & dbh class
11	No. of growing-stock trees on timberland by species group & dbh class
12	Net volume of growing stock on timberland by species group & dbh class
13	Net cu. ft. volume in sawlog portion of sawtimber trees on timberland by species group & dbh class
14	Net volume of sawtimber on timberland by species group & dbh class
15	Net volume of growing stock & sawtimber on timberland by county & major species group
16	Net volume of timber on timberland by class of timber & major species group
17	Net volume of live trees & growing stock on timberland by ownership class & major species group
18	Net ann. growth of growing stock & sawtimber on timberland by county & major species group
19	Ave. ann. removals of growing stock & sawtimber on timberland by county & major species group
20	Net ann. growth & removals of growing stock on timberland by species group
21	Net ann. growth & removals of sawtimber on timberland by species group
22	Ann. mortality of growing stock & sawtimber on timberland by species group
23	Net ann. growth & removals of growing stock on timberland by owner class & major species group
24	Net ann. growth & removals of sawtimber on timberland by owner class & major species group
25	Net volume of sawtimber on timberland by species group & tree grade

EWDB STRUCTURE

The EWDB can be thought of as having a simple hierarchical structure. This structure corresponds to FIA plot collection and data processing methods and contains data at three levels. The highest level is the county, followed by the plot level, and finally, the tree level.

Data is stored as one of three record types (10-county record, 20-plot record, and 30-tree record), each with its own format. The format, a detailed description of each data item, and a complete description of the EWDB structure are presented in a separate EWDB users manual (Hansen, et al., in prep.). Here I briefly describe the structure and format of each record. These records or data tables contain data for one level of the hierarchy.

The EWDB consists of separate sequential files for each state made up of the three record types and grouped by county. Each county record is followed by the plot record for the first plot in that county. Following each plot record, are the tree records associated with the plot. Not all plots have tree records, such as plots on nonforested land. Table 2 shows the name of each data item or attribute in these records, the FORTRAN format for reading the item, the columns occupied and a brief description. This table also contains a column indicating the key data items which uniquely identify an occurrence of a data record.

Table 2. Data items in the FIA EWDB.

County record				
Item	Format	Columns*	Description	Key
STATE	I2	3-4	State code	X
UNIT	I1	5	Survey unit number	
COUNTY	I3	6-8	County code	X
CTYNAM	A28	13-40	County name	
STNAME	A2	41-42	State name (postal abbreviation)	
CYCLE	I2	43-44	Survey cycle number	
DATE	I4	45-48	Date of the survey (year)	

Plot record				
Item	Format	Columns*	Description	Key
STATE	I2	3-4	State code	X
UNIT	I1	5	Survey unit number	
COUNTY	I3	6-8	County code	X
PLTNUM	I4	9-12	Plot number	X
OWNER	I2	13-14	Ownership code	
TYPCUR	I2	15-16	Forest type - current inventory	
TYPOLD	I2	17-18	Forest type - last inventory	
STDAGE	I3	19-21	Stand age (years)	
STDSIZE	I1	22	Stand size class	
STORCUR	I1	23	Stand origin - current inventory	
STOROLD	I1	24	Stand origin - last inventory	
SITECL	I1	25	Site class	
SI	I2	26-27	Site index (feet)	
SIAGE	I2	28-29	Site index base age (years)	
ADFOR	I3	30-32	National Forest code	
GLUCUR	I2	33-34	Land use - current inventory	
GLUOLD	I2	35-36	Land use - last inventory	
BA	I3	37-39	Basal area (square feet)	
SLOPE	I2	40-41	Slope (percent)	
ASPECT	I3	42-44	Aspect (degrees)	
PHYSIO	I1	45	Physiographic class	
TREATOP	I2	46-47	Treatment opportunity class	
INHIBPC	I2	48-49	Percent occupied by inhibiting vegitation	
NONSTPC	I2	50-51	Percent nonstocked	
GRSTKPC	I3	52-54	Percent stocked with growing stock trees	
ALSTKPC	I3	55-57	Percent stocked with live trees	
REMPER	F3.1	58-60	Remeasurement period (years)	
EXPACR	I5	61-65	Area expansion factor (acres)	
EXPVOL	I5	66-70	Volume expansion factor (acres)	
EXPGRO	I5	71-75	Growth expansion factor (acres)	
EXPMOR	I5	76-80	Mortality expansion factor (acres)	
EXPREM	I5	81-85	Removals expansion factor (acres)	
LONG	F7.1	86-92	Longitude (seconds)	
LAT	F7.1	93-99	Latitude (seconds)	

* Columns 1-2 contain the record number (10=county, 20=plot, 30=tree)

(Table 2 is continued on the next page)

Table 2 (continued). Data items in the FIA EWDB.

Tree record				
Item	Format	Columns*	Description	Key
STATE	I2	3-4	State code	X
UNIT	I1	5	Survey unit number	
COUNTY	I3	6-8	County code	X
PLTNUM	I4	9-12	Plot number	X
POINT	I2	13-14	Point number	X
TREE	I2	15-16	Tree number	X
STATUS	I1	17	Tree status (live, dead or cut)	X
SPP	I3	18-20	Species	
SPPGRP	I2	21-22	Species group	
DBHCUR	F3.1	23-25	DBH - current inventory (inches)	
DBHOLD	F3.1	26-28	DBH - last inventory (inches)	
TGRADE	I1	29	Tree grade	
TCLASS	I1	30	Tree class (growing stock, rough or rotten)	
CRATIO	I1	31	Crown ratio	
CRCLS	I1	32	Crown class	
DAMAGE	I2	33-34	Damage class	
VOLFAC	F8.4	35-42	Volume expansion factor (trees/acre)	
MORTFAC	F8.4	43-50	Mortality expansion factor (trees/acre)	
REMVFAC	F8.4	51-58	Removals expansion factor (trees/acre)	
NETCFVL	F8.4	59-66	Net cubic foot volume	
NETCFSL	F8.4	67-74	Net cubic foot volume in sawlog portion	
NETBFVL	F8.4	75-82	Net board foot volume	
NETCFGR	F8.4	83-90	Net annual cubic foot growth	
NETBFGR	F8.4	91-98	Net annual board foot growth	
TOTBIO	I6	99-104	Total gross biomass (green pounds)	
MERBIO	I6	105-110	Merchantable biomass (green pounds)	

* Columns 1-2 contain the record number (10=county, 20=plot, 30=tree)

USING THE EWDB

The flat file structure of the EWDB was designed to be used easily with most database management systems, statistical packages, or other data summary software. It is not dependent on a particular hardware or software system, but the system must accept input in the form of a flat ASCII or EBCDIC file up to 110 characters wide. EWDB files have been loaded and analyzed by a number of commercial software packages, including SAS, ORACLE, dBASE III, RBASE, INGRES, SIR, SPSS and Foxbase, on hardware ranging from portable microcomputers to large mainframes. Database management systems that support hierarchical data structures as well as those based on the relational model can easily process EWDB files. The EWDB users manual should provide the user with the information needed to input an EWDB file into their processing system. To assist users of EWDB files and to provide a benchmark or checkpoint for comparison, the FIA projects will provide a set of the 25 core tables with any EWDB file.

EWDB file users will require the procedure or algorithm used to compute various tree level data and expand it to population level estimates. Table 3 summarizes how this is done for many commonly requested data elements. Using Table 3, the proper way to compute the total number of all live trees on timberland in a state from the EWDB file for that state would be to multiply the value in VOLFAC on the tree record by the value in EXPVOL on the matching plot record, and sum this product (VOLFAC*EXPVOL) over every tree record where STATUS is equal to one and GLUCUR on the plot record is equal to 20. If you were interested in knowing the current number of live trees over 5.0 inches DBH, the process would be the same, with the added restriction that only tree records with DBHCUR over 5.0 are included in the summation.

Table 3. Algorithms to compute various tree level information expanded to population level estimates. Each piece of informantion is computed by summing the corresponding function of data items over all trees that meet the test requirements.

Information	Function	Test
Current number of trees on timberland		
All live	VOLFAC*EXPVOL	GLUCUR=20 & STATUS=1
Growing stock	VOLFAC*EXPVOL	GLUCUR=20 & STATUS=1 & TCLASS=2
Salvable dead	VOLFAC*EXPVOL	GLUCUR=20 & STATUS=4
Current volume on timberland		
All live merch.	NETCFVL*VOLFAC*EXPVOL	GLUCUR=20 & STATUS=1
Growing stock	NETCFVL*VOLFAC*EXPVOL	GLUCUR=20 & STATUS=1 & TCLASS=2
Sawtimber	NETBFVL*VOLFAC*EXPVOL	GLUCUR=20 & STATUS=1 & TCLASS=2
Rough trees	NETCFVL*VOLFAC*EXPVOL	GLUCUR=20 & STATUS=1 & TCLASS=3
Rotten trees	NETCFVL*VOLFAC*EXPVOL	GLUCUR=20 & STATUS=1 & TCLASS=4
Net annual growth on timberland		
Growing stock	NETCFGR*VOLFAC*EXPGRO	GLUCUR=20 OR GLUOLD=2O
Sawtimber	NETBFGR*VOLFAC*EXPGRO	GLUCUR=20 OR GLUOLD=2O
Annual mortality on timberland		
Growing stock	NETCFVL*MORTFAC*EXPMOR	GLUCUR=20 OR GLUOLD=2O & TCLASS=2
Sawtimber	NETBFVL*MORTFAC*EXPMOR	GLUCUR=20 OR GLUOLD=2O & TCLASS=2
Annual removals from timberland		
Growing stock	NETCFVL*REMVFAC*EXPREM	TCLASS=2
Sawtimber	NETBFVL*REMVFAC*EXPREM	TCLASS=2
Biomass on timberland		
All live total	TOTBIO*VOLFAC*EXPVOL	GLUCUR=20 & STATUS=1
All live merch.	MERBIO*VOLFAC*EXPVOL	GLUCUR=20 & STATUS=1

Those familiar with the relational data model and the standard Structured Query Language (SQL) database language available in many database management systems will find it easy to load EWDB files into one of these systems and to retrieve information from a loaded database. EWDB files are written in normal form as defined by the relational model. The two retrievals described in the last paragraph are relatively easy to perform using SQL. The SQL command to obtain the total number of all live trees on timberland in a state from a database containing data from just that state would be:

```
select  sum(VOLFAC*EXPVOL)           from  PLOT, TREE
 where  PLOT.COUNTY = TREE.COUNTY     and  PLOT.PLTNUM = TREE.PLTNUM
   and  TREE.STATUS =1                and  PLOT.GLUCUR = 20
```

and the SQL command to retrieve the same information for trees over 5.0 inches dbh would be:

```
select  sum(VOLFAC*EXPVOL)           from  PLOT, TREE
 where  PLOT.COUNTY = TREE.COUNTY     and  PLOT.PLTNUM = TREE.PLTNUM
   and  TREE.STATUS =1                and  PLOT.GLUCUR = 20
   and  DBHCUR > 5.0.
```

It is suggested that EWDB users accessing with SQL begin with a SQL query designed to retrieve a known quantity in one of the core tables such as the first example in the last paragraph. Once you verify that your basic retrieval is working correctly you can modify it to retrieve only the data of interest to you. For example you may be interested in the volume of select white oak sawtimber in trees over 20 inches dbh in a four county area. Begin by testing a retrieval to get the total sawtimber volume in the four county area using the following SQL query:

```
select   sum(NETBFVL*VOLFAC*EXPVOL)   from   PLOT, TREE
where    PLOT.COUNTY = TREE.COUNTY     and    PLOT.PLTNUM = TREE.PLTNUM
  and    TREE.STATUS =1                and    PLOT.GLUCUR = 20
  and    TCLASS=2
  and   (PLOT.COUNTY = C1 or PLOT.COUNTY = C2 or PLOT.COUNTY = C3 or
         PLOT.COUNTY = C4)
```

where C1, C2, C3 and C4 are the county codes of the four counties of interest. Once you verify that this query is working by comparing the total volume retrieved to core table 15 you can modify it the query to get only select white oak trees (SPPGRP = 10) at least 20 inches in dbh. This SQL query would look like this:

```
select   sum(NETBFVL*VOLFAC*EXPVOL)   from   PLOT, TREE
where    PLOT.COUNTY = TREE.COUNTY     and    PLOT.PLTNUM = TREE.PLTNUM
  and    TREE.STATUS =1                and    PLOT.GLUCUR = 20
  and    TCLASS=2
  and   (PLOT.COUNTY = C1 or PLOT.COUNTY = C2 or PLOT.COUNTY = C3 or
         PLOT.COUNTY = C4)             and    TREE.SPPGRP = 10
  and    DBHCUR >= 20.0.
```

For those interested in obtaining additional information on the relational data model and SQL there are many references available including Date (1986 and 1987), Cronin (1989), Cardenas (1979) and Everest (1986). Many of the database management systems that support SQL have excellent user manuals and tutorial guides for their implementations of SQL. Many of these systems also support enhancements that go beyond standard SQL, allowing additional flexibility and formating, of output or that take advantage of other features of the database management system.

OBTAINING EWDB FILES

EWDB files for a state must be obtained from the FIA project responsible for the inventory of that state. Currently, files can be obtained from all projects on standard 9 track tapes (1600 or 6250 BPI, ASCII or EBCDIC). There is a charge of $500 per State to cover the cost of producing files and maintaining the system. Table 4 shows the status of each State in the EWDB system, the date of the last inventory, the date when the next inventory is scheduled to be completed, and the approximate size of the EWDB files and the availability of these files. These figures are based on projected schedules at the time this report was written. For current information on a particular State or how to order an EWDB file, contact the individual FIA project shown in figure 1.

Some projects may be able to provide EWDB files on a medium more easily read by microcomputers. Because technology and equipment for data storage and transportation is rapidly changing, some projects may be able to provide these files on diskettes, data cartridges or laser disks for an additional charge. Specific information is available from the individual projects.

Table 4. Status and size of each EWDB file.

State	Current inv. date	EWDB files		Next scheduled inventory
		Date avail.	Size (MB)	
- - - - - - - - North Central Region - - - - - - - -				
IA	1974	12/90	2	1989
IL	1985	12/89	6	1996
IN	1986	09/89	10	1997
KS	1981	12/89	5	1992
MI	1980	12/89	25	1993
MN	1977	12/89	23	1991
MO	1989	09/90	12	1999
NE	1983	12/89	2	1992
ND	1980	12/89	3	1992
SD	1980	12/89	2	1992
WI	1983	12/89	21	1995
- - - - - - - - - - Southern Region - - - - - - - - - -				
AL	1982	09/89	10	1990
AR	1988	09/89	14	1994
LA	1984	09/89	7	1991
MS	1987	09/89	13	1994
OK	1986	09/89	3	1992
TN	1980	09/89	8	1989
TX	1986	09/89	9	1992

* Information not available at time of publication

State	Current inv. date	EWDB files		Next scheduled inventory
		Date avail.	Size (MB)	
- - - - - - - - - Northeast Region - - - - - - - - -				
CT	1984	*	<1	1996
DE	1986	*	<1	1997
KY	1987	*	5	1998
ME	1982	*	6	1994
MD	1985	*	3	1997
MA	1984	*	<1	1996
NH	1983	*	2	1995
NJ	1986	*	<1	1996
NY	1980	*	6	1992
OH	1979	*	4	1990
PA	1978	*	6	1990
RI	1984	*	<1	1996
VT	1983	*	2	1995
WV	1988	*	4	2000
- - - - - - - - - Southeast Region - - - - - - - - -				
FL	1987	09/89	10	1993
GA	1989	09/89	17	1994
NC	1984	09/89	12	1990
SC	1986	09/89	9	1992
VA	1986	09/89	10	1991

REFERENCES

Cardenas, A. F., 1979 Data Base Management Systems, Allyn and Bacon, Boston, 519 pp.

Cronin, D. J., 1989 Mastering ORACLE, Featuring ORACLE's SQL Standard, Hayden, Indianapolis, 543 pp.

Date, C. J., 1986 An Introduction to Database Systems, Addison-Wesley, Reading, Mass., 639 pp.

Date, C. J., 1987 A Guide to the SQL Standard, Addison-Wesley, Reading, Mass., 639 pp.

Everest, G. C., 1986 Database Management, Objectives, System Functions, and Administration, McGraw-Hill, New York, 816 pp.

Hansen, M. H., Frieswyk, T. S., Glover, J. F., and Kelly, J. F., (in prep), A Users Guide to the Eastwide Forest Inventory Database, USDA, Forest Service, Washington, DC.

Access To Forest Inventory Data For The Southeast

Raymond M. Sheffield
Research Forester
Forest Inventory and Analysis
Southeastern Forest Experiment Station
Asheville, NC 28802

ABSTRACT

Forest inventory data in the Southeast are highly accessible through publications, Forest Information Retrieval, special data sets, customized programming, and graphical displays. Specific information needs of forest resource data users are met in an efficient and cost effective manner through consultation and selection of appropriate dissemination methods.

INTRODUCTION

The Southeastern Forest Experiment Station's Forest Inventory and Analysis Research Work Unit (SE-FIA) has a special interest in the topic being addressed during this session. Making the forest inventory data we collect available to a wide range of users has been considered a priority--a "work ethic" if you will--for the past 20 to 30 years. Our goal has been, and continues to be, efficient and accurate use of the data by all clients who have valid uses. These objectives require substantial consultation with clients--interaction that is mutually beneficial. Users need the contact to become familiar with our data and to determine their most appropriate uses of those data. FIA staff members benefit from the interaction by keeping in touch with the real world as we learn of problems and issues our clients face. Ultimately, improvements in the inventory process result.

TYPES OF DATA REQUESTS

A wide range of tabular data is published in Resource Bulletins in the Southeast. The information is presented for individual counties, for regions within each State, and for entire States. Forest resource data requirements for most of our users are satisfied by these published sources. When more specific data are required, several options are available: (1) Forest Information Retrieval (FIR), (2) special data sets, (3) customized programming with output, and (4) graphical displays.

Forest Information Retrieval (FIR)

Over the last two decades, a software package and associated data called Forest Information Retrieval (FIR) have been used thousands of times to provide customized forest resource data for the Southeast (McClure 1972). Now, as in the past, approximately 70 percent of special requests for data are filled by FIR. Even with the increased availability of FIA data in other forms, demand for the tabular data provided by FIR has held firm. We have found that many of the people requesting FIA data do not have the expertise or equipment to manipulate their own data. In addition, many users simply do not want to spend their time in this manner.

The FIR system provides up to 81 tables of forest resource information, including data describing ownership, forest type and species composition, timber volume, tree sizes, and rates of timber growth, harvest and mortality. These tables can be compiled for any geographic area within the Southeast a user may delineate. An area can be defined by a circle around a known point, a group of whole counties, or an irregular boundary specified by the user. Built into the FIR system is the flexibility to compile resource statistics for only that portion of the total resource the user is interested in. For instance, one can compile tables for only certain ownerships, certain types of stands such as pine plantations, or stands with some minimum size and volume per acre.

The FIR system can be used to provide consistent information for up to three consecutive surveys in most Southeastern States. FIR tables are usually provided to users very rapidly. The time required can be as short as 30 minutes; key statistics are sometimes phoned to the person requesting the data. Usually, the turn-around averages 2 to 3 days depending on project workload. Costs also vary depending upon the size of the area of interest, the number of tables obtained, and other variables such as screening and the method of defining the geographic area. The cost for obtaining all 81 tables for a group of 50 whole counties, without screening, would be $99.

A users guide is available upon request (U.S. Department of Agriculture 1988). It describes the capabilities and uses of the FIR system and provides a cost algorithm.

Special Data Sets

A growing number of forest resource data users are requesting actual data sets instead of printed tables. Many people now have personal computers or access to larger computers and they want to be able to analyze the data themselves. A major advantage of having the data residing on their own computers is the ability to answer an array of questions that may change from day to day. Currently, approximately 20 percent of all requests for customized forest resource information are filled by providing actual data to the user. Data provided may range from county-level statistics to detailed tree-level data. Large data sets usually are provided on magnetic tape; smaller files often are sent on floppy disks.

Special data sets require more time to compile than standard FIR tables. The time needed depends on the complexity of the data requested and our workload at the time. Response time ranges from 1 week for simple requests to 6 weeks for more complicated ones. Costs also vary widely. Most requests for specialized data sets cost $300 or more.

Customized Programming

Clients sometimes have very specific needs that cannot be met through FIR. An FIA staff member consults with these users to define the needs. Together, they agree on types of information needed to answer key questions and on a format for displaying the data. Customized programming usually is designed to provide tables, but customized data sets also are provided.

Customized programming is involved in about 5 percent of all requests. Again, response time and cost vary with the complexity of the required software development, the nature of the data to be used, and the size of the geographic area involved. Response time is similar to that for data sets--1 to 6 weeks. Costs are also similar, with $300 representing a minimum in most cases. Typical requests involving special software development cost between $500 and $1000.

Graphical Output

Graphical displays of FIA data are an effective means of communicating results to clients. This response medium is growing in importance, but still makes up less than 5 percent of all special requests. Types of output are very user specific and require a considerable amount of staff interaction with the client. Currently, most examples of graphical output are map-based products showing resource distributions.

Accurate estimates of cost and response time for graphical output are not possible at this time because these products are still in the developmental stage.

FUTURE CHANGES

In the foreseeable future, we anticipate only one major change in how users access SE-FIA data. We expect to provide more graphical displays to users. Graphs will become a standard component of FIR output, enhancing the interpretation of the tabular data now provided. A variety of products based on a geographic information system will become available as software is perfected and put into a production mode.

Notably absent from our future plans is a switch to interactive use of FIA data bases. Existing systems make FIA data in the Southeast highly accessible with short response times. The current system is user friendly. Therefore, few users have voiced an interest in having direct access to our data bases.

REFERENCES

McClure, Joe P., 1972: Customized forest information retrieval. Forest Farmer, vol. 31, no. 10, pp. 6-7.

U.S. Department of Agriculture, Forest Service, 1988: A user's guide to forest information retrieval (FIR). U.S. Dep. Agric. For. Serv., Asheville, NC: Southeastern Forest Experiment Station, 99 pp.

Interactive Access to Forest Inventory Data for the South Central United States

William H. McWilliams
Research Forester

USDA, Forest Service
Southern Forest Experiment Station
Box 906, Starkville, MS 39759

ABSTRACT

On-line access to USDA, Forest Service successive forest inventory data for the South Central United States is provided by two computer systems. The Easy Access to Forest Inventory and Analysis Tables program (EZTAB) produces a set of tables for specific geographic areas. The Interactive Graphics and Retrieval System (INGRES) is a database management system that supplies access to inventory data at three levels of aggregation.

INTRODUCTION

Multi-resource inventories in seven states of the South Central region and Puerto Rico are conducted periodically by the USDA, Forest Service, Southern Forest Experiment Station, Forest Inventory and Analysis unit (SO-FIA or FIA). The states include Alabama, Arkansas, Louisiana, Mississippi, Oklahoma, Tennessee, and Texas. A highly dynamic forest resource and a broad based user group require that the results of the inventories be made available in as rapid, efficient, and flexible a manner as possible. Traditional publications are being supplemented by providing users with direct access to forest inventory data via a table generation program and a database management system[1].

TABLE GENERATION

The EZTAB program is a menu driven system that is available to users interested in tabular information for specific geographic areas. The program compiles tables for any configuration of inventory plots, be it a circle, polygon, or combination of counties. Very often, a request is made for plots within a given radius of a location -- for example, within 50 miles of Selma, Alabama. The EZTAB program provides a set of 43 tables, including 25 "core" tables that are standardized in publications of the eastern FIA units. Subsets of tables can be selected as desired. If needed, SO-FIA personnel will generate EZTAB tables for users who forego direct computer access. A fee is charged for this service.

---

[1] USDA, Forest Service. 1987: Interactive data access manual. Unpublished manual. Starkville, MS: U.S. Department of Agriculture Forest Service, Southern Forest Experiment Station. On file with: Forest Inventory and Analysis, Box 906, Starkville, MS, 39759.

DATABASE MANAGEMENT

INGRES is a relational database management system that expands the ability to access forest inventory data. The INGRES system accommodates retrieval of information at three levels: county, plot, and tree. The system includes 27 county data sets for the South Central region, as well as 14 plot-level and 2 tree-level data sets for each state inventoried by SO-FIA. Data sets are accessible for the two most recent inventories of each state. Each county and plot-level data set provides access to a particular type of information, such as timberland area, inventory volume, growth, removals, biomass, and nontimber resources. Tree data broaden the flexibility for making retrievals, but additional computations are required to obtain population values. Database queries may be done using the INGRES Query Language (QUEL) or the widely distributed Structured Query language (SQL). More advanced options for INGRES processing include interfacing with high-level languages, report writing, and graphics.

DATA ACCESS

Interactive access to forest inventory data requires only a computer terminal, a telephone, and a modem. Users are charged a fee based on the amount of computer resources used. The fee also includes the cost of a 2-day workshop that covers EZTAB, INGRES, and the SO-FIA computer system. Workshops are scheduled periodically to meet user demand.

FURTHER INFORMATION

The SO-FIA computer currently supports a variety of users that include forest industries, consultants, universities, and state agencies. For information about interactive access to forest invnentory data, prospective users may contact:

Data Request Coordinator
USDA, Forest Service
Forest Inventory and Analysis
Box 906, Starkville, MS, 39759
(601) 324-1611

North Central Forest Inventory Data on a Microcomputer

Jerold T. Hahn
and
Mark H. Hansen
Research Foresters
USDA, Forest Service
North Central Forest Experiment Station
St. Paul, Minnesota, 55108

ABSTRACT

Microcomputer versions of the North Central Forest Inventory data base are now or will soon be available for all 11 states (Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, South Dakota, and Wisconsin) inventoried by the Forest Inventory and Analysis (FIA) work unit at North Central Forest Experiment Station. These data bases operate on XT or AT level IBM PC compatible microcomputers under the SIR data base management system.

INTRODUCTION

How much aspen is within 50 miles of Grand Rapids, Minnesota? How many acres of 20-to-40-year-old red pine plantations are in Michigan? What is the current volume of white oak sawtimber in southern Indiana? These questions can be answered by using data from periodic inventories conducted by the Forest Inventory and Analysis (FIA) Work Unit of the North Central Forest Experiment Station, USDA Forest Service, in St. Paul, MN. Providing answers to questions like these is our job, and we are constantly searching for ways to provide clients all the information possible in a versatile and efficient form.

Recent advances in data base management software and the widespread use of microcomputers have given microcomputer users the same data retrieval capabilities once available only to mainframe computer users.

NORTH CENTRAL FOREST INVENTORY DATA BASE

Station researchers began to develop the North Central Forest Inventory Data Base (NCFIDB) in 1978 (Murphy 1981). Since the 1983 Wisconsin inventory, we have used the NCFIDB (Hahn and Hansen 1985) to store and process all FIA data and to answer special requests for information. This data base is maintained on a Data General MV/20,000 computer and operates under the SIR Data Base Management System (DBMS) (Robinson et al. 1980).

The data base contains an interactive, menu-driven interface that allows a user to generate forest resource tables. Users with some computer programming experience, a knowledge of Forest Service inventory and the SIR DBMS, have more flexibility in the kinds of retrievals they can make by using the SIR 4GL and/or SQL queries. We are currently developing procedures for outside users to access these data bases directly on our system.

Several outside users have accessed our data base directly in the past when it was being maintained on a mainframe timeshare system at the University of Minnesota. However, time on the mainframe was costly and users had to learn the operating system. Many users had their own computer systems (mainframe, mini, and/or micro), thus making it difficult for them to justify spending money on an outside computer.

Therefore, we began searching for a system that would allow outside users to access our data on their own machines. These efforts are continuing since we do not anticipate that there will be extra resources available on our own system to meet the needs of outside users.

MICROCOMPUTER DATA BASE MANAGEMENT SYSTEMS

The growing popularity and power of microcomputers has made processing and analyzing large data sets on these systems more common. DBMS software programs allow users to perform a series of coordinated data handling functions, including describing the data and storing it in accordance with this description, extracting information from the data, and protecting the data from unauthorized or accidental modification. These systems make the data easier to use and more portable.

dBASE III

The first attempt to access FIA data on a microcomputer was a cooperative effort of the Forest Service, the North Central Computer Institute, and the Department of Forest Resources, University of Minnesota. In this effort we used the popular dBASE III (Ashton-Tate Company 1984) DBMS. We developed a dBASE III data file structure and wrote procedures to download data from the mainframe SIR data base to dBASE III data files. In addition, "command files" or manipulation and reporting programs for dBASE III were written to facilitate user access and to provide a starting point for further analysis. This method of accessing FIA data with a microcomputer is described by Belli et al. (1986).

One problem with the dBASE III system was that the many data retrieval programs written by FIA personnel for the mainframe data base are in SIR retrieval language, not dBASE III command files. To fully support the dBASE III system would require maintaining two sets of retrieval programs (SIR and dBASE III) and constant checking to ensure both systems were producing the same results when changes were made to the system. Therefore, only a few example dBASE III command files were written, and users were limited to working from them to obtain retrievals.

SIR

Recently the SIR DBMS (version 2.3) was released for use on IBM PC compatible microcomputers. Earlier versions of PC SIR did not support all the features of SIR that are used by the NCFIDB on the mainframe. But now we can directly download portions of the NCFIDB to PC's. The download procedure transfers any part of the data base, such as a portion of the data, data structure, retrieval procedures, or any other procedures used by the system.

The micro and mainframe versions of the data base operate identically, and PC SIR users have the same flexibility in working with the data as mainframe users. Essentially, the only requirements for the PC user are the SIR DBMS software, a copy of the relevant section of the NCFIDB, and a microcomputer capable of operating the software.

The minimum computer requirements for using the SIR DBMS system are:

An IBM PC/XT, PC/AT or 100 percent compatible microcomputer
640 kilobytes or more of available memory
One hard disk with at least 7.5 megabytes available
8087/80287 mathematical coprocessor chip
One double-density floppy drive
DOS version 2.0 or higher.

To efficiently use the NCFIDB on a PC, a larger (40+ megabytes) hard disk is highly recommended. Data bases for areas covering several counties in heavily forested regions can easily exceed 20 megabytes. A 20 megabyte, 8-inch Bernoulli box is also very useful. Because of the large size of the data base files, it is easier to transport these files on Bernoulli cartridges than on floppy disks. We currently do not have the facility to write to storage devices other than floppy disks or Bernoulli cartridges from our PC's. The SIR DBMS can be purchased directly from SIR Inc. (707 Lake Cook Road, Suite 120, Deerfield, IL 60015).

On the mainframe, the NCFIDB is a single "logical" data base, stored as 11 physical units, each covering a state (fig. 1, Table 1). By a single logical data base, we mean that throughout the entire 11-state region the exact same data structure is used, and all retrieval programs will run on every state. The data base has been broken into these 11 portions for efficiency and because storage is limited on the mainframe.

For use on the PC, the NCFIDB had to be broken into even smaller portions (fig. 2, Table 2). SIR data bases are limited to the 32 megabytes per disk drive partition imposed by the DOS operating system. We have tried to limit the size of each data base to approximately 20 megabytes for the loaded or operational form of the data base. When a data base is to be transported, it is unloaded into a compressed format. The unloaded form column of Table 2 shows how much space is needed to transport these data bases. Table 2 also shows the size of each PC SIR data base in its loaded form. For example, although the three data bases that make up the Michigan Lower Peninsula (MICH3N, MICH3S, AND MICH4) can all be transported together on a single 20-megabyte Bernoulli cartridge, only one data base could be loaded and used on a 20-megabyte drive. Users with a large disk drive can merge several data bases together into a single data base on their machines. The Wisconsin Department of Natural Resources has made a data base for the entire state of Wisconsin on an 80-megabyte drive on a 386 machine.

If you would like copies of any of the PC SIR data, contact us (612-649-5139). You must provide enough blank 20-megabyte 8-inch Bernoulli cartridges or 5 1/4-inch disks to hold the data base requested. You will be charged a fee for any data requested, to defray the expenses we incur in providing this service. Data bases are shipped as SIR unload files and include not only the data, but also a complete data structure and a number of data retrieval and manipulation programs.

Access by the Casual User

Once you have the required hardware, loaded SIR on your hard drive, and obtained a data base from us, you can retrieve forest inventory data without learning any special programming language or FIA methods. Using the interactive retrieval program that comes with each data base, you will soon be able to obtain forest resource tables for any area in the North Central Region. Tables can be constructed for a county or a group of counties, a circular area of any size and location, or an irregular area. Tables can be generated for all forest stands in an area or only for those of a selected forest type.

To obtain a particular resource table for any area, simply follow a few simple instructions to load the desired data base on your hard disk, then respond to questions, and choose from menus. You can choose from a list of available tables (42 tables are currently available, and more tables are being added). The table is then constructed and automatically saved as an ASCII file for later viewing, printing, and/or processing by some other software such as a spreadsheet or word processor. Then repeat the process for another table and/or a different area. Table 3, for example, is for a user interested in the forest resource of a four-county area in north central Minnesota. The time required to produce a table will depend on the particular table desired, the size of the area selected, and the speed of your machine. Table 3 took 25 minutes to make on my AT&T 6310. Three minutes of this was the time it took to enter the data base, request the appropriate table, and define the four county area; the remaining 22 minutes was the processor time to make the table. Processor time will vary depending on table selected, method of selecting stand(s) (circle, counties, polygon, forest type, or all stands), and the speed of your machine.

In the future, we plan to expand this program so that users with a basic knowledge of Forest Service inventory methods can also use the same program to construct files of inventory data for analysis by any statistical package, or for input to their own programs written in any language. By responding to questions and making menu selections, the user will select the items to be included in the file. Information about what items are available and definitions of each item will be available interactively from the program.

If you would like a personal demonstration of the PC SIR and the NCFIDB, contact us or stop by the North Central Station in St. Paul.

Table 1. North Central Forest Inventory Data Bases on the Mainframe.

Data base	State	Date of inventory	Number of: Total plots	Forest plots	Trees
MICHST	Michigan	1980	14,490	8,371	315,538
WISCST	Wisconsin	1983	15,920	8,144	288,066
MINNST	Minnesota	1977	36,530	10,543	287,993
NDAKOTA	North Dakota	1980	18,214	604	5,339
SDAKOTA	South Dakota (eastern)	1980	23,206	518	2,928
NEBRASK	Nebraska	1983	14,449	535	5,004
KANSAS	Kansas	1981	23,705	1,807	36,306
MISSOUR	Missouri	1972	10,611	2,992	95,101
IOWA	Iowa	1974	12,607	756	16,276
ILLINI	Illinois	1985	10,957	1,327	42,496
INDIANA	Indiana	1986	11,440	2,829	94,171

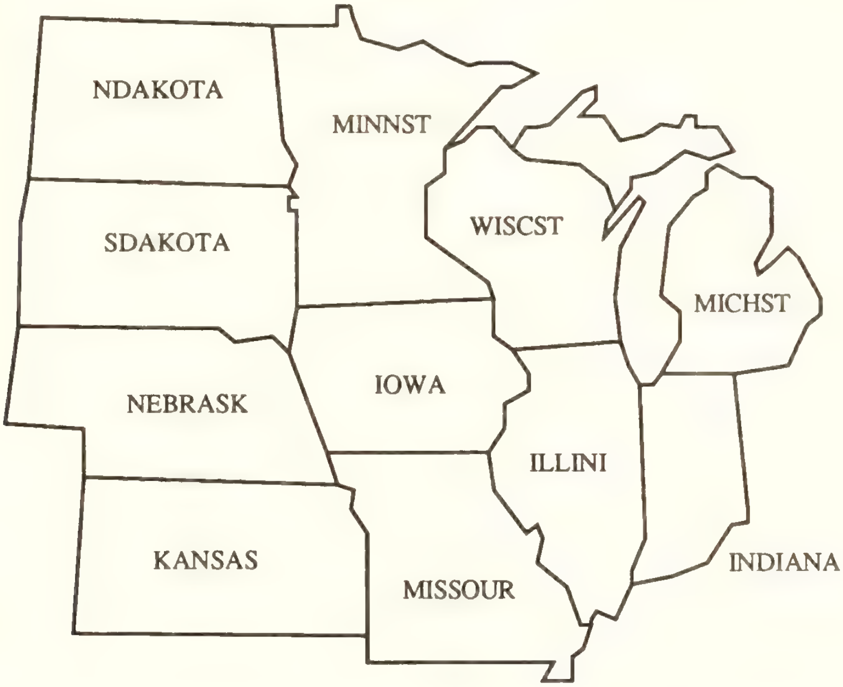


Figure 1. North Central Forest Inventory Data Bases on the Mainframe

Table 2. North Central Forest Inventory Data Bases currently available on the PC.

Data base	Number of Counties	Total plots	Forest plots	Trees	Megabytes of storage Loaded[a]	Unloaded[b]
MICH1E	3	1,253	992	37,537	12.2	6.0
MICH1W	4	1,487	1,182	46,698	6.1	8.3
MICH2E	4	1,760	1,520	62,115	20.2	11.6
MICH2W	4	1,219	1,015	43,394	15.9	11.2
MICH3N	15	2,134	1,470	53,445	14.2	7.7
MICH3S	18	2,610	1,441	47,597	13.0	7.1
MICH4	35	4,027	751	24,752	6.8	4.8
WISC1N	5	1,636	1,315	56,211	16.1	8.5
WISC1S	5	937	796	32,362	11.0	5.9
WISC2N	7	2,525	2,164	76,776	23.0	11.8
WISC2S	5	1,656	1,058	32,184	10.1	5.6
WISC3	13	3,195	1,539	51,455	16.9	8.8
WISC4	14	2,455	820	25,535	6.5	5.0
WISC5	23	2,516	452	13,543	5.8	3.3
IND1	14	1,803	581	19,385	6.7	4.8
IND2	17	2,113	1,107	40,820	11.8	6.5
IND3	9	802	367	11,310	4.5	3.0
IND4	52	6,722	774	22,656	9.6	5.1
MINN1E	2	852	647	21,331	10.9	5.8
MINN1W	3	4,268	3,286	89,098	24.3	12.5
MINN2N	6	4,332	1,861	46,669	13.9	8.2
MINN2S	6	4,263	2,576	81,277	23.6	12.1
MINN3	28	8,830	1,716	38,049	10.2	6.2
MINN4	42	13,985	457	11,579	7.1	5.3

[a] Form the data base takes when in use on the user's hard disk.

[b] Form used to ship the data base on Bernoulli cartridge or floppy disks.

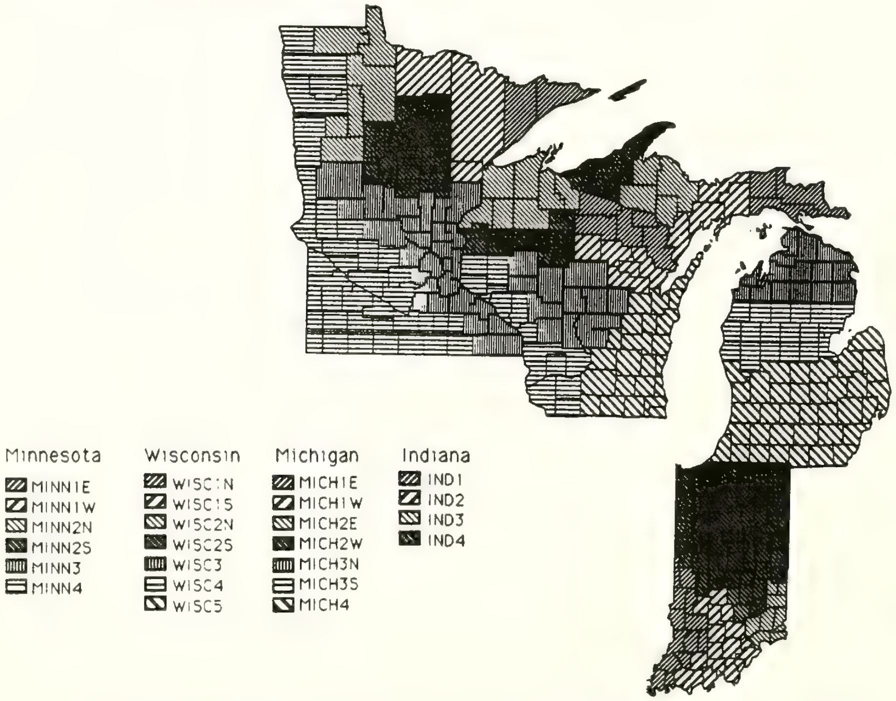


Figure 2. North Central Forest Inventory Data Bases currently available on the PC

Table 3. A custom table produced by the interactive retrieval program from data for four counties[a].

Area of timberland by forest type and stand size class,
Selected counties, Minnesota, 1977
(In thousand acres)

Forest type	Stand size class				
	All classes	Saw-timber	Pole timber	Sapling & Seedling	Non-stocked
	Thousand acres				
Jack pine	174.8	49.2	104.7	20.9	*
Red pine	104.2	28.0	16.8	59.4	*
White pine	9.0	2.5	6.5	*	*
Balsam fir	25.8	4.3	12.9	8.6	*
White spruce	4.7	1.4	1.9	1.4	*
Black spruce	28.6	1.1	15.4	12.1	*
Northern white-cedar	30.8	8.9	16.5	5.4	*
Tamarack	53.4	3.2	30.4	19.8	*
Oak-hickory	167.5	46.6	97.0	23.8	*
Elm-ash-soft maple	52.7	10.1	31.2	11.4	*
Maple-basswood	82.0	30.2	47.1	4.7	*
Aspen	851.5	173.1	488.4	190.0	*
Paper birch	137.8	13.6	114.0	10.2	*
Balsam poplar	24.7	8.7	10.3	5.7	*
Nonstocked	7.9	*	*	*	7.9
All types	1,755.3	380.9	993.1	373.4	7.9

[a] This run retrieved data for the following counties

County		Number of plots retrieved
Code	Name	
204	Cass	503
206	Crow Wing	262
207	Hubbard	307
212	Wadena	74
	All counties	1146

Access by the Programming User

With a fundamental knowledge of FIA methods and the SIR DBMS, a user with computer programming skills has greater flexibility in the kinds of retrievals that can be made from the database and in the format of the tables produced. For example, such a user could retrieve area and volume information for a specific geographic area. Suppose that the only stands of interest are those whose volume is composed of 20 percent aspen, and that the desired output is a table showing area by forest type and stand age, along with tables showing total volume (all species) and volume of aspen by forest type and owner. Suppose further that the geographic area is defined as all land within a 50-km radius of a proposed mill site.

Such a task is relatively easy for the programming user and can be done using a program less than 50 lines long. However, this type of access does require the user to have a much higher level of knowledge, and data retrieved under these circumstances are more likely to be misinterpreted. For these reasons, we urge anyone seeking information from the database to try first to obtain it through the interactive retrieval program. Usually this program will satisfy the user's information needs; only rarely will he/she need to try the programming approach. For information on survey methods used and details of the data base structure, the programming user is referred to "Data Bases for Forest Inventory in the North-Central Region" (Hahn and Hansen 1985).

We hope that the PC version of the NCFIDB will help meet the information needs of many organizations. We are trying to make FIA data more accessible to the user, and hope this effort will be useful to many of you. We welcome suggestions on how to improve the data base and retrieval programs for your use.

REFERENCES

Ashton Tate Co. 1984. dBASE III, Version 1.1. User Manual. 365 p.

Belli, K. L.; Ek, A. R.; Hansen, M. H.; and Hahn, J. T. 1986. A region-wide forestry database for microcomputers. NCCI Software J. 113 p.

Hahn, J. T., and M. H. Hansen. 1985. Data bases for forest inventory in the North Central Region. USDA For. Serv. Gen. Tech. Rep. NC-101. 57 p.

Murphy, D. L. 1981. Database management: a forest inventory application. USDA For. Serv. Gen. Tech. Rep. NC-67. 9 p.

Robinson, B. N., G. D. Anderson, E. Cohen, W. F. Gazdzik, L. C. Karpel, A. H. Miller, and J. R. Stein. 1980. SIR Scientific Information Retrieval user's manual--version 2. SIR Inc., Evanston, Il. 475 p.

Successive Forest Inventories: Basic SPR Approach For Two Measurements[1]

Tiberius Cunia
Professor of Statistics and Operations Research
State University of New York College of Environmental Science and Forestry
Syracuse NY 13210 USA

ABSTRACT

The basic SPR approach to two successive measurements of the same population is described. Basic formulae are given for the estimation of (i) the means of the first and second measurement, (ii) the net mean change from the first to the second measurement, and (iii) the precision of these estimators. A numerical example is also given to illustrate the application of the SPR approach and the efficiency of the SPR estimators.

INTRODUCTION

Most of the time, the populations one samples are dynamic. For example, the trees growing in a forest area constitute a dynamic population that changes continuously with time; trees are continuously added by birth processes (trees growing from seeds) or lost by death processes (trees that die or are removed by harvesting) and the trees change their size and characteristics all the time.

If one is interested in the state of the forest population at a given point in time, and he has little, if any interest in the rates of change, a simple survey by the usual sampling methods found in standard textbooks would suffice. Similarly, when one is interested in the changes occurring in the population over a given period of time, the usual sampling designs applied to sample units measured at the beginning and the end of that period would also be adequate. But if one is continuouly interested in both, the current values and the rates of change, these sampling designs are no longer efficient.

It is the objective of this discussion to present an approach which increases the efficiency of the survey sampling when a population is monitored over time for both, current values and changes with time. We shall consider the specific case of a population, in our case the trees growing in a given forest area, which is measured on successive occasions. On each occasion, a standard sampling design of the usual type is applied and the current values of the parameter of interest are being estimated. The rates of change would normally be estimated by the differences in the current values from one to the next occasion.

It is obvious that some of the sample units of one occasion may also be sample units of another occasion, by chance or by design. The approach we shall discuss here considers the way by which the samples of various occasions are related to each other. Because (i) some sample units are specifically selected to be part of the samples of several occasions ("permanent" sample units) and (ii) some "old" units of one occasion are dropped from the sample of the next occasion and replaced by "new" units, ("temporary" sample units) it seems reasonable to call this approach "sampling with partial replacement" or SPR.

To simplify the discussion, we shall consider here the case of two successive measurements. We shall then suggest, in a companion paper, Cunia (1989a), the ways by which the SPR approach can be extended to cover the case where (i) auxiliary variables may be used to define more efficient estimators, (ii) several parameters are simultaneously estimated and (iii) the population is measured on more than two occasions. Because it is common for forest inventory systems to use biomass (volume) regression functions, and the error of these regressions is not taken into account when the error of biomass (volume) estimates is evaluated, we shall also describe, in another paper, Cunia (1989b), an approach that combines the regression theory with the theory of SPR as presently applied in forest inventory (which implicity assumes that the attributes of the sample units are "measured" not "estimated" by regression).

---

[1] Paper based on a set of class notes for a graduate course in Forest Sampling.

To better see how the efficiency of estimators may be affected by the way by which the sample units of two consecutive measurements are "chained", let us consider a specific, hypothetical sampling situation. Assume that a given forest area has been inventoried on two successive occasions, say, five years apart, and that each time, the sample units (plots) were selected by simple random sampling. If (i) x and y denote the volume per plot as measured on the first and second occasion, and (ii) we have the same sample size n on both occasions, let us define the basic statistics $\bar{x}$ and $\bar{y}$ as the sample mean volume per plot on the first and second occasion respectively and S_{xx} and S_{yy} the corresponding variances.

Let us assume first that the units of the second occasion were selected independently of the units of the sample of the first occasion. Then, elementary statistical theory tells us that (i) $\bar{x}$ and $\bar{y}$ have variances estimated by S_{xx}/n and S_{yy}/n respectively and (ii) the net change from the first to the second occasion is estimated by $(\bar{y}-\bar{x})$ with variance estimated by $(S_{xx}/n + S_{yy}/n)$. Let us make now the assumption that the same sample plots were measured on both occasions. Then, the variables x and y measured on the same plot are correlated, with r = correlation coefficient and $S_{xy} = r\sqrt{S_{xx}\, S_{yy}}$ = covariance, and (i) $\bar{x}$ and $\bar{y}$ have the same variance as before, but (ii) the variance of the net change $(\bar{y}-\bar{x})$ is now estimated by $(S_{xx}/n + S_{yy}/n - 2S_{xy}/n)$. As the covariance of x and y is relatively large, the variance of the net change in this last case is much smaller.

To have an idea of the magnitude of the errors, let us use the values $n = 100$, $\bar{x} = 200$, $\bar{y} = 220$, $S_{xx} = 10000$, $S_{yy} = 14400$, $r = .95$ and $S_{xy} = r\sqrt{S_{xx}\, S_{yy}} = 11400$, all, very reasonable values for many forest areas. Then, the estimate of the net change is $(\bar{y}-\bar{x}) = 20$ and the estimate of its variance is (i) $(S_{xx}/n + S_{yy}/n) = 244$ for the first sampling design where the samples are statistically independent and (ii) $(S_{xx}/n + S_{yy}/n - 2S_{xy}/n) = 16$ for the second sampling design where the same sample plots are measured on both occasions. As the reader can see, the variance of the estimates of the current mean values did not change (same variances of $S_{xx}/n = 100$ and $S_{yy}/n = 144$) but the variance of the net change, under the second sampling design, where the same units are measured on both occasions has decreased considerably. Assuming that the sampling costs are the same, the efficiency ratio is 144/16 = 9, that is, the efficiency of the second design with respect to the estimation of the net change has increased to 900 percent (compared to the first sampling design).

It can be shown that the SPR approach, which replaces a part of the first occasion sample by a new set of plots will (i) reduce slightly the efficiency of the net change estimate (when compared to the second design where all the plots are measured on both occasions) but (ii) increase dramatically the efficiency of the current mean values. In addition, the SPR approach can do more: it can be applied with the sample size that would yield exactly the desired precision in each of the estimates of mean values and net change, for minimum cost of sampling. With the other two designs, the sample size is determined by the estimator whose desired precision requires the largest sample size.

BASIC SPR DESIGN AND ESTIMATORS FOR TWO MEASUREMENT OCCASIONS

The general sampling with partial replacement design for measuring a population on two successive occasions consists of (i) a first occasion sample of $n_1 = (u + m)$ units measured for values x and (ii) a second occasion sample of $n_2 = (m + n)$ units (m of which are part of the n_1 units of the sample of the first occasion) measured for values y. Here, u stands for "unmatched", m for "matched" and n for "new" sample units.

The above definition being too general, we shall make the following simplifying assumptions: (i) the basic procedure for selecting the units at each of the two occasions is Simple Random Sampling without Replacement, (ii) the same N population units are present at both measurement occasions (that is, there are no birth or death processes to change the population), and (iii) the sample sizes n_1 and n_2 are small with respect to the population size N, so that the effect of the finite population correction factor is negligibly small.

To define the estimators of μ_x = arithmetic mean at measurement 1 time, μ_y = arithmetic mean at measurement 2 time and $\mu_d = \mu_y - \mu_x$ = arithmetic mean of the change from the first to the second measurement time, when σ_{xx}, σ_{yy}, σ_{xy} and ρ denote respectively variance of x, variance of y, covariance of x and y and correlation coefficient, we shall denote the sample values as x_k, $k = 1, 2, \ldots, u, u + 1, \ldots, u + m = n_1$ and y_k, $k = u + 1, u + 2, \ldots, u + m, u + m + 1, \ldots, u + m + n$. That is, we shall write the sample values x as

$$x_1, x_2, \ldots, x_u, x_{u+1}, x_{u+2}, \ldots, x_{u+m}$$

and the sample values y as

$$y_{u+1}, y_{u+2}, \ldots, y_{u+m}, y_{u+m+1}, \ldots, y_{u+m+n}$$

The basic statistics we shall use are $\overline{x}_u$ = mean of values x for the unmatched u units, $\overline{x}_m$ = means of values x for the matched m units, $\overline{y}_m$ = mean of values y for the matched m units, $\overline{y}_n$ = mean of values y for the new n units, and the variances and covariance S_{xx}, S_{yy} and S_{xy}, calculated from the matched m units only. We shall also use the statistics $\overline{x} = (u\overline{x}_u + m\overline{x}_m)/(u + m)$ = mean of all values x of the (u + m) units measured on the first occasion and a similar statistic $\overline{y}$ = mean of all values y of the (m + n) units measured on the second occasion.

Using these statistics we can define the estimators: (i) $\overline{x}_u$, $\overline{x}_m$ and $\overline{x}$ of μ_x; (ii) $\overline{y}_m$, $\overline{y}_n$ and $\overline{y}$ of μ_y; and (iii) $(\overline{y}_m - \overline{x}_m) = \overline{d}_m$ and $(\overline{y} - \overline{x}) = \overline{d}$ of μ_d. The several estimators of the same parameter are not equal in precision; among those mentioned above, $\overline{x}$, $\overline{y}$ and oddly enough (because it is based on a part of the data) $\overline{d}_m$ are the best. These estimators will be known here as Non-SPR estimators, to distinguish them from the SPR estimators defined below which, in all cases are better.

The SPR estimators of μ_y, μ_x and $\mu_d = (\mu_y - \mu_x)$ and their variances are defined as follows

(1) $\hat{y} = \overline{y}_n + a_y(\overline{x}_u - \overline{x}_m) + c_y(\overline{y}_n - \overline{y}_m)$ = SPR estimator of μ_y

where

$a_y = umb_{y \cdot x}/D$; $c_y = -m(u + m)/D$; $D = (u + m)(m + n) - unr^2$

$b_{y \cdot x} = S_{xy}/S_{xx}$ = slope of the least squares line of y on x

$r^2 = S^2_{xy}/S_{xx} S_{yy}$ = square of correlation coefficient

(2) $S_{\hat{y}\hat{y}} = (1 + c_y) S_{yy}/n$ = estimator of the variance of $\hat{y}$

Because the sampling design is symmetrical with repect to x and y, the formulae of $\hat{x}$ and S_{xx} follow immediately.

(3) $\hat{x} = \overline{x}_u + a_x(\overline{x}_u - \overline{x}_m) + c_x(\overline{y}_n - \overline{y}_m)$ = SPR estimator of μ_x

where

$a_x = -m(m + n)/D$; $c_x = mnb_{x \cdot y}/D$

$b_{x \cdot y} = S_{xy}/S_{yy}$ = slope of the least square line of x on y

(4) $S_{\hat{x}\hat{x}} = (1 + a_x) S_{xx}/u$ = estimator of the variance of $\hat{x}$

Finally, the SPR estimator of $\mu_d = (\mu_y - \mu_x)$ is

(5) $\hat{d} = (\overline{y}_n - \overline{x}_u) + a_d(\overline{x}_u - \overline{x}_m) + c_d(\overline{y}_n - \overline{y}_m) = \hat{y} - \hat{x}$ = SPR estimator of μ_d

where

$a_d = a_y - a_x = (m(m + n) + umb_{y \cdot x})/D$

$c_d = c_y - c_x = -(m(u + m) + mnb_{x \cdot y})/D$

(6) $S_{\hat{d}\hat{d}} = (1 - a_d) S_{xx}/u + (1 + c_d) S_{yy}/n$ = estimate of the variance of $\hat{d}$

To estimate the $(1 - \alpha)$ confidence intervals and the totals for the N population elements, one can use the well known formulae

estimator $\pm\, t_{\alpha/2}\sqrt{\text{variance of estimator}}$

$\hat{T} = (N)$ (estimator of the mean) = estimator of the total

and $S_{\hat{T}\hat{T}} = (N^2)$ (variance of estimator) = variance of $\hat{T}$, respectively,

where $t_{\alpha/2}$ value is read from a table of t-distribution with (m - 1) degrees of freedom.

Remarks:

(1) At the first measurement occasion one cannot calculate $\hat{x}$; the best estimator at that time is $\bar{x}$; But on the second occasion one may revise the estimate of μ_x by using $\hat{x}$. The efficiency of $\hat{x}$ relative to that of $\bar{x}$ can be expressed by the ratio

$$S_{\bar{x}\bar{x}}/S_{\hat{x}\hat{x}} = u/(u + m)(1 + a_x) > 1, \text{ since } a_x > 0$$

(2) The SPR estimators can be expressed in many equivalent ways. They all consist of a sample mean whose expected value is the parameter to estimate and two adjustments for the differences between the means of the samples of u, m and n units, whose expected values are zero. For example, for the case of $\hat{y}$, we have the sample means $\bar{y}_n$, $\bar{y}_m$ or $\bar{y}$ (whose expected value is μ_y) and the adjustments for the differences $(\bar{x}_u - \bar{x}_m)$, $(\bar{x}_u - \bar{x})$, etc.. and $(\bar{y}_m - \bar{y}_n)$, $(\bar{y} - \bar{y}_m)$, etc. whose expected value is zero. Of particular importance is the expression

$$\hat{y} = \bar{y} + a\,(\bar{x}_u - \bar{x}_m) + c\,(\bar{y} - \bar{y}_n)$$

where $a = a_y$ and $c = unr^2/D$, with a_y, D, u, n and r^2 already defined. One can view this SPR estimator $\hat{y}$ as being equal to the best Non-SPR estimator $\bar{y}$ (the mean of all measured values y) adjusted for the difference between $\bar{y}$ and $\bar{y}_n$ and the difference between $\bar{x}_m$ and $\bar{x}_u$. The variance of the SPR estimator expressed in this form can be written as

$$S_{\hat{y}\hat{y}} = S_{yy}/(m + n) - (S_{yy}/(m + n))(um\, r^2/D)$$

a difference between two components; a first component being $S_{yy}/(m + n)$, the variance of $\bar{y}$, the best Non-SPR estimator and a second component, $um\, r^2\, S_{yy}/D(m + n)$, which shows by what amount the variance of $\bar{y}$ is reduced because of the use of the information provided by the unmatched u units. Note that (i) the second component becomes zero whenever $u = 0$, $m = 0$ or $r^2 = 0$, (ii) $S_{\hat{y}\hat{y}} < S_{\bar{y}\bar{y}}$ whenever $r^2 > 0$ (and $u \neq 0$, $m \neq 0$) and (iii) $S_{\hat{y}\hat{y}} = S_{yy}/(m + n + u)$, whenever $r^2 = 1$, that is $\bar{y}$ has the same precision as if all $(u + m + n)$ were actually measured.

(3) Another useful interpretation of $\bar{y}$ is based on the split of the SPR sample of $(u + m + n)$ units into two statistically independent subsamples; the first subsample of $(u + m)$ units with the sample values $x_1, x_2, \ldots, x_{u+m}$ and $y_{u+1}, y_{u+2}, \ldots, y_{u+m}$ and the second subsample of the n new units with the values $y_{u+m+1}, y_{u+m+2}, \ldots, y_{u+m+n}$. The data of the first subsample is used to define a double sampling estimator of μ_y, say

$$\bar{y}_d = \bar{y}_m - b_{y\cdot x}\,(\bar{x}_m - \bar{x})$$

with variance estimated as approximately equal to

$$S_{\bar{y}_d\bar{y}_d} = S_{yy}\,(1-r^2)/m + r^2 S_{yy}/(u + m)$$

and the data of the second subsample provides the estimator $\bar{y}_n$ with variance $S_{\bar{y}_n\bar{y}_n} = S_{yy}/n$. Then, it can be shown that the SPR estimator $\hat{y}$ is identically equal to the weighted mean $\bar{y}_w = w_d\,\bar{y}_d + w_n\,\bar{y}_n$ where $w_d = S_{\bar{y}_n\bar{y}_n}/(S_{\bar{y}_d\bar{y}_d} + S_{\bar{y}_n\bar{y}_n})$ and $w_n = S_{\bar{y}_d\bar{y}_d}/(S_{\bar{y}_d\bar{y}_d} + S_{\bar{y}_n\bar{y}_n})$. Furthermore, alegebraic manipulations of the variance formula show that we have, as we should,

$$S_{\bar{y}_w\bar{y}_w} = w^2{}_d\, S_{\bar{y}_d\bar{y}_d} + w^2{}_n\, S_{\bar{y}_n\bar{y}_n} = S_{\hat{y}\hat{y}}$$

This new interpretation of the SPR estimator leads to two useful extensions. One extension is the formula which takes into account the effect of the finite population correction factor, since one can show that an approximate formula for the variance of $\bar{y}_d$ and an exact one for the variance of $\bar{y}_n$ are

$$S_{\bar{y}_d\bar{y}_d} = (u\, S_{yy\cdot x}/(u + m))(1/m + (\bar{x}-\bar{x}_m)^2/(m - 1)S_{xx}) + (N - u - m)\, Syy/N(u + m)$$

and

$$S_{\bar{y}_n\bar{y}_n} = (N - n)\, S_{yy}/Nn$$

respectively where N is the population size. The second extension allows one to select the units within subsample 1 and those of subsample 2 by methods other than simple random sampling. If these sampling methods are such that one can find good estimators $S_{y_dy_d}$ and $S_{y_ny_n}$, one can define the SPR estimators for sampling designs which are much more complex than the one assumed by our basic formulae.

(4) We shall not consider here the problem of determining the future sample size. It is first too complex as (i) one must take into account the costs of sampling various types of units, c_u, c_m, c_n, etc. and (ii) one must satisfy conditions about desired precision of the various estimators at various points in time. The solution to the general problem requires dynamic or convex programming techniques, although sometimes, in simple cases, one may use graphical methods. It is questionable, however, whether it is worthwhile spending the required time and effort, when the best solution may be quite sensitive to the assumed values of the various parameters that enter into the model.

AN ILLUSTRATIVE NUMERICAL EXAMPLE

To illustrate how the SPR estimators are calculated and to realize some of the efficiency obtained, we apply this approach to a set of data from 81 plots listed in Table 1. The variable x and y denote the plot volume in cunits (hundreds of cubic feet), expressed on a per acre basis, as measured on two successive occasions four years apart, say in 1960 and 1964.

Table 1 - Volumes per acre (cunits) of 81 randomly selected plots measured on two successive occasions in 1960 and 1964.

Plot	Volume/Acre		Plot	Volume/Acre		Plot	Volume/acre	
Number	first (x)	second (y)	Number	first(x)	second(y)	Number	first(x)	second(y)
1	13.1	- -	28	5.7	7.3	55	4.6	4.7
2	8.4	- -	29	16.5	17.7	56	12.1	10.0
3	29.5	- -	30	11.2	10.3	57	24.1	26.7
4	22.7	- -	31	11.6	12.5	58	19.3	21.6
5	16.7	- -	32	14.8	17.5	59	14.8	16.2
6	13.2	- -	33	51.6	52.8	60	- -	5.5
7	1.4	- -	34	16.4	18.3	61	- -	14.4
8	21.3	- -	35	33.4	35.7	62	- -	21.7
9	37.6	- -	36	10.6	13.8	63	- -	18.5
10	7.1	- -	37	14.9	17.0	64	- -	11.0
11	19.6	- -	38	5.1	7.2	65	- -	8.7
12	23.0	- -	39	11.2	11.1	66	- -	4.4
13	21.0	- -	40	.8	1.6	67	- -	20.8
14	22.2	- -	41	14.0	16.5	68	- -	27.3
15	25.1	- -	42	16.7	19.1	69	- -	20.5
16	20.7	- -	43	35.1	35.0	70	- -	10.0
17	17.8	- -	44	4.9	7.5	71	- -	7.8
18	12.1	- -	45	14.3	16.2	72	- -	34.9
19	2.1	3.1	46	22.7	26.6	73	- -	19.6
20	2.9	4.5	47	15.8	17.4	74	- -	14.1
21	6.8	8.5	48	12.6	15.7	75	- -	3.4
22	8.9	11.1	49	24.6	25.4	76	- -	22.7
23	1.8	2.2	50	19.2	21.0	77	- -	18.4
24	11.0	12.6	51	8.8	11.0	78	- -	24.1
25	6.4	8.0	52	14.9	16.6	79	- -	13.8
26	31.1	35.6	53	14.4	14.4	80	- -	28.8
27	12.8	14.4	54	13.8	15.8	81	- -	22.9

We start with the calculation of the basic sums, sums of squares and sum of crossproducts. $u = 18$, $m = 41$, $n = 22$, $\Sigma_u\, x = 332.5$, $\Sigma_m\, x = 594.3$, $\Sigma_m\, y = 660.2$, $\Sigma_n\, y = 373.3$, $\Sigma_u\, x^2 = 7369.41$, $\Sigma_m\, x^2 = 12621.89$, $\Sigma_m\, y^2 = 14858.34$, $\Sigma_m\, xy = 13657.04$, $\Sigma_n\, y^2 = 7817.31$

We continue with the calculation of the basic statistics $\bar{x}_u = 18.47222222$, $\bar{x}_m = 14.49512195$, $\bar{y}_m = 16.10243902$, $\bar{y}_n = 16.96818182$, $S_{xx} = 100.1859756$, $S_{xy} = 102.1840122$, $S_{yy} = 105.6877439$, $b_{y \cdot x} = 1.019943276$, $b_{x \cdot y} = .9668482685$, $r^2 = b_{y \cdot x} b_{x \cdot y} = .9861303907$

This leads immediately to the calculation of the SPR coefficients

$a_y = .2262798333$, $a_x = -.7764935904$, $a_d = a_y - a_x = 1.002773424$

$c_y = -.7271924101$, $c_x = .2621671847$, $c_d = c_y - c_x = -.9893595947$

and that of the SPR estimators and their variances.

(1) $\hat{y} = 16.96818182 + .89993758 - .62956159 = 17.23855781$

= estimate of μ_y, the 1960 mean volume per acre (second measurement)

where .89993758 and -.62956159 are the adjustments (made to $\bar{y}_n = 16.96818182$) due to the differences ($\bar{x}_u - \bar{x}_m$) and ($\bar{y}_n - \bar{y}_m$) respectively. As the variance of $\hat{y}$ is estimated as

$S_{\hat{y}\hat{y}} = (1 + c_y)\, S_{yy}/n = 1.310564486$,

the 95 percent confidence limits of μ_y (using $t = 2$) are

(17.24 ± 2.29) cunits per acre

Note that the non-SPR estimator is

$\bar{y} = \Sigma y/(m + n) = 1033.5/63 = 16.40476190$

with a variance estimated as

$S_{yy}/(m + n) = 105.6877439/63 = 1.677583237$

Consequently the efficiency ration of $\hat{y}$ is

$S_{\bar{y}\bar{y}}/S_{\hat{y}\hat{y}} = n/(n + m)(1 + c_y) = 22/(63)(.2728075899) = 1.280046311$

It may be of interest to show the interpretation of $\hat{y}$ as a weighted mean of a "double sampling" and a "simple random sampling" estimator, $\hat{y} = w_d\bar{y}_d + w_n\bar{y}_n$.

$\bar{y}_d = 17.33998988$, $\bar{y}_n = 16.96818182$, $S_{\bar{y}_d\bar{y}_d} = 1.802225200$,

$S_{\bar{y}_n\bar{y}_n} = 4.803988359$, $w_d = .7271924100$, $w_n = .2728075900$ and, thus,

$\hat{y} = w_d\bar{y}_d + w_n\bar{y}_n = 17.23855782$ and $S_{yy} = 1.310564487$

(2) $\bar{x} = \Sigma x/(u + m) = (332.5 + 594.3)/59 = 15.70847458$

= non-SPR estimate of μ_x, the 1960 mean volume per acre (first measurement) based on the data available in 1960.

$\hat{x} = 15.61099868$ = the SPR estimate (an update of $\bar{x}$) of μ_x, the 1960 mean volume per acre based on all sample data available in 1964

As the variance of $\hat{x}$ is estimated by

$S_{\hat{x}\hat{x}} = (1 + a_x)\, S_{xx}/u = 1.244011539$

the 95 percent confidence limits of μ_x are evaluated as

(15.61 ± 2.23) cunits per acre

Note that the efficiency ratio of $\hat{x}$ relative to $\bar{x}$ is

$S_{\bar{x}\bar{x}}/S_{\hat{x}\hat{x}} = u/(u + m)(1 + a_x) = 1.365$

(3) $\hat{d} = \hat{y} - \hat{x} = 17.23855781 - 15.61099868 = 1.62755911$

= estimate of the net change in the volume per acre between 1960 and 1964

$S_{\hat{d}\hat{d}} = (1 - a_d)\, S_{xx}/u + (1 + c_d)\, S_{yy}/n = .035678917$

and the 95 percent confidence limits are

(1.628 ± .378) cunits per acre

Note that the best, non-SPR estimator is

$\bar{d}m = \bar{y}_m - \bar{x}_m = 1.60731707$

with variance estimated as

$S_{\bar{d}m\bar{d}m} = (S_{yy} - 2\, S_{xy} + S_{xx})/m = .03672427073$

The efficiency ratio of the SPR estimator $\hat{d}$ relative to $\bar{y}$ is

$S_{\bar{d}m\bar{d}m}/S_{\hat{d}\hat{d}} = .03672427073/.035678917 = 1.03$

a negligible value.

On the other hand, one may feel that the best non-SPR estimator would be the difference between the best estimators of μ_y and μ_x, namely $\bar{y}$ and $\bar{x}$. Then $\bar{d} = \bar{y} - \bar{x} = .69628732$ with variance

$$S_{\bar{d}\bar{d}} = S_{yy}/(m + n) - 2m\, S_{xy}/(u + m)(m + n) + S_{xx}\,(u + m) = 1.121389388$$

As the reader can verify, $\bar{d}$ is a very poor estimator.

SPECIAL CASES

When one of the subsamples of u, m or n units are missing, several formulae yield values of 0/0 which, in order to be evaluated, require taking their limit. For the reader unaccustomed to the process of evaluation of ratios 0/0, we shall give below, the formulae in the form in which they can be used.

Special Case 1: m = 0 - Independent samples on each occasion.

$\hat{y} = \hat{y}_n$, $\hat{x} = \bar{x}_u$, $\hat{d} = \bar{y}_n - \bar{x}_u$

$S_{\hat{y}\hat{y}} = S_{yy}/n$, $S_{\hat{x}\hat{x}} = S_{xx}/u$ and $S_{\hat{d}\hat{d}} = S_{yy}/n + S_{xx}/u$

Special Case 2: u = n = 0 - Matched units only

$\hat{y} = \bar{y}_m$, $\hat{x} = x_m$, $\hat{d} = \bar{y}_m - \bar{x}_m$

$S_{\hat{y}\hat{y}} = S_{yy}/m$, $S_{\hat{x}\hat{x}} = S_{xx}/m$ and $S_{\hat{d}\hat{d}} = (S_{yy} - 2S_{xy} + S_{xx})/m$

Special Case 3: u = 0 - All units of first occasion measured on the second occasion

$\hat{y} = \bar{y}$ = simple arithmetic mean of the (m + n) values y

$\hat{x} = \bar{x}_d = \bar{x}_m - b_{x \cdot y}\,(\bar{y}_m - \bar{y})$ = double sampling estimator of μ_x

$\hat{d} = \hat{y} - \hat{x} = (\bar{y} - \bar{x}_m) + b_{x \cdot y}\,(\bar{y}_m - \bar{y})$

$S_{\hat{y}\hat{y}} = S_{yy}/(m + n)$, $S_{\hat{x}\hat{x}} = S_{xx}\,(1 - r^2)/m + r^2\, S_{xx}/(m + n)$

and

$S_{\hat{d}\hat{d}} = (S_{yy} - 2S_{xy} + S_{xx})/(m + n) + n(1 - r^2)\, S_{xx}/n(m + n)$

Special Case 4: n = 0 ; No new units are added at the second occasion

$\hat{y} = \bar{y}_d = \bar{y}_m - b_{y \cdot x}\,(\bar{x}_m - \bar{x})$ = double sampling estimator of μ_y

$\hat{x} = \bar{x}$ = simple arithmetic mean of the (u + m) values x

$\hat{d} = \hat{y} - \hat{x} = (\bar{y}_m - \bar{x}) - b_{y \cdot x}\,(\bar{x}_m - \bar{x})$

$S_{\hat{y}\hat{y}} = S_{yy}(1 - r^2)/m + r^2 S_{yy}/(u + m)$, $S_{\hat{x}\hat{x}} = S_{xx}/(u + m)$

and

$S_{\hat{d}\hat{d}} = (S_{yy} - 2S_{xy} + S_{xx})/(u + m) + u(1 - r^2)\, S_{yy}/u(u + m)$

REFERENCES

Cunia, T. (1989a): Successive forest inventories: extensions of the basic approach. Proceedings of the "State of the Art Methodology of Forest Inventory" Symposium, July 31-August 6, 1989, SUNY College of Environmental Science and Forestry, Syracuse NY 13210.

Cunia T. (1989b): Successive forest inventories: error of biomass regressions. Part 2: CFI with partial replacement of plots. Proceedings of the "State of the Art Methodology of Forest Inventory" Symposium, July 31-August 6, 1989, SUNY College of Environmental Science and Forestry, Syracuse NY 13210.

Successive Forest Inventories: Extensions of the Basic Approach[1]

Tiberius Cunia
Professor of Statistics and Operations Research
State University of New York College of Environmental Science and Forestry
Syracuse NY 13210

ABSTRACT

The methodology of SPR is extended from the basic case of the measurement of a population on two occasions, where each sample unit is measured for one variable of interest only (variable x on the first and variable y on the second occasion) and the estimator is calculated for one parameter at a time to the more complex cases where (i) sample units are measured for p variables x at the first and q variables y at the second occasions and all these variables are used in the estimation process, (ii) the population is measured on three or more occasions and (iii) several parameters are estimated simultaneously. The extension requires the presentation of the SPR methodology in a matrix form.

INTRODUCTION

To extend the basic approach of SPR on two successive occasions described by Cunia (1989a) to the case of (i) sample units measured for auxiliary variables in addition to the basic variables of interest x and y, (ii) three or more successive occasions and (iii) multivariate estimation, we shall present the basic SPR estimators of Cunia (1989a) in a matrix form. We shall assume that the reader is familiar with (i) matrix algebra and (ii) the notation used by Cunia (1989a). The matrices will be denoted here by the notation [ ] and the inverse and transpose matrices by $[\ ]^{-1}$ and [ ]' respectively. The sample data will consist of variables x measured on the (u + m) sample units of the first occasion and variables y measured on the (m + n) sample units of the second occasion. Note that m represents the number of sample units measured on both occasions. The summary statistics are (i) the sample means $\bar{x}_u$, $\bar{x}_m$, $\bar{y}_m$ and $\bar{y}_n$, of the subsamples of u "unmatched" units measured for x alone, m "matched" units measured for both x and y and n "new" units measured for y alone, and (ii) the sample variances and covariances S_{xx}, S_{yy} and S_{xy} calculated from the subsample of m "matched" units only. The SPR estimators of μ_x, μ_y and $\mu_d = (\mu_y - \mu_x)$ are denoted by $\hat{x}$, $\hat{y}$ and $\hat{d}$ respectively. For more details on these statistics the reader is referred to Cunia (1989a).

THE MATRIX APPROACH

We start by defining the following vectors and matrices of sample data associated wtih SPR estimator y of μ_y.

$[Q_y] = [y_n]$ = vector (of order 1) of unmatched (new) sample mean

$[P] = \begin{bmatrix} \bar{x}_u - \bar{x}_m \\ \bar{y}_n - \bar{y}_m \end{bmatrix}$ = vector of sample differences between the means of matched and unmatched units

$[K_y] = [S_{yy}/n]$ = sample covariance matrix of Q_y]

$[H_y] = \begin{bmatrix} 0 \\ S_{yy}/n \end{bmatrix}$ = sample covariance matrix of [P] and $[Q_y]$

---

[1] Paper based on a set of class notes for a graduate course in Forest Sampling.

$$[G] = \begin{bmatrix} S_{xx}\,(1/u + 1/m) & S_{xy}\,(1/m) \\ S_{xy}\,(1/m) & S_{yy}\,(1/m + 1/n) \end{bmatrix} = \text{sample covariance matrix of } [P]$$

Then, the SPR estimators can be written as

(1) $\hat{y} = [Q_y] + [A_y]'[P] = [Q_y] - [H_y]'[G]^{-1}[P]$ = SPR estimator of μ_y

where

$$[A_y] = \begin{bmatrix} a_y \\ \\ c_y \end{bmatrix} = -[G]^{-1}\,[H_y] = \text{vector of SPR coeffficients}$$

and

(2) $S_{\hat{y}\hat{y}} = [K_y] + [A_y]'\,[H_y] = [K_y] - [H_y]'\,[G]^{-1}\,[H_y]$ = estimate of the variance of $\hat{y}$

Because of symmetry, the SPR formulae for $\hat{x}$ can be similarly written, with a similar notation as

(1) $\hat{x} = [Q_x] + [A_x]'\,[P] = [Q_x] - [H_x]'\,[G]^{-1}\,[P]$ = SPR estimator of

where

$$[Q_x] = [\bar{x}_u]\,,\quad [H_x] = \begin{bmatrix} S_{xx}/u \\ \\ 0 \end{bmatrix} \text{and } [A] = -[G]^{-1}\,[H_x]$$

and

[2] $S_{\hat{x}\hat{x}} = [K_x] + [A_x]'\,[H_x] = [K_x] - [H_x]'\,[G]^{-1}\,[H_x]$ = estimate of the variance of $\hat{x}$

where

$[K_x] = [S_{xx}/u]$ = sample covariance matrix of $[Q_x]$

Extended to the SPR estimation of $\mu_d = (\mu_y - \mu_x)$, we can similarly define and write

(1) $\hat{d} = [Q_d] + [A_d]'\,[P] = [Q_d] - [H_d]'[G]^{-1}[P] = \big(\text{SPR estimator of } \mu_d = (\mu_y - \mu_x)\big) = \hat{y} - \hat{x}$

where

$[Q_d] = [\bar{y}_n - \bar{x}_u] = [Q_y] - [Q_x]$

$[K_d] = [S_{yy}/n + S_{xx}/u] = [K_y] + [K_x]$ = sample covariance matrix of $[Q_d]$

$$[H_d] = \begin{bmatrix} -S_{xx}/u \\ \\ S_{yy}/n \end{bmatrix} = \begin{bmatrix} 0 \\ \\ S_{yy}/n \end{bmatrix} - \begin{bmatrix} S_{xx}/u \\ \\ 0 \end{bmatrix} = [H_y] - [H_x]$$

= sample covariance matrix of [P] and $[Q_d]$

and

$$[A_d] = \begin{bmatrix} a_d \\ \\ c_d \end{bmatrix} = -[G]^{-1}\,[H_d] = [A_y] - [A_x] = \text{vector of SPR coefficients}$$

and

(2) $S_{\hat{d}\hat{d}} = [K_d] + [A_d]'\,[H_d] = [K_d] - [H_d]'\,[G]^{-1}\,H_d]$ = estimate of the variance of d.

We can now combine all three SPR estimators into one set of formulae by defining, in specific terms,

$$[\hat{\mu}] = \begin{bmatrix} \hat{y} \\ \hat{x} \\ \hat{d} \end{bmatrix} = \text{SPR estimator of } [\mu] \text{ defined as} \begin{bmatrix} \mu_y \\ \mu_x \\ \mu_d \end{bmatrix}$$

$$[Q] = \begin{bmatrix} \bar{y}_n \\ \bar{x}_m \\ \bar{y}_n - \bar{x}_u \end{bmatrix} = \text{vector of ummatched sample means}$$

$$[P] = \begin{bmatrix} \bar{x}_u - \bar{x}_m \\ \\ \bar{y}_n - \bar{y}_m \end{bmatrix} = \text{vector of differences between matched and unmatched sample means}$$

$$[K] = \begin{bmatrix} S_{yy}/n & 0 & S_{yy}/n \\ 0 & S_{xx}/u & -S_{xx}/u \\ S_{yy}/n & -S_{xx}/u & (S_{yy}/n + S_{xx}/u) \end{bmatrix} = \text{sample covariance matrix of } [Q]$$

$$[H] = \begin{bmatrix} 0 & S_{xx}/u & -S_{xx}/u \\ \\ S_{yy}/n & 0 & S_{yy}/n \end{bmatrix} = \text{sample covariance matrix of } [P] \text{ and } [Q]$$

$$[G] = \begin{bmatrix} S_{xx}\ (1/u + 1/m & S_{xy}\ (1/m) \\ \\ S_{xy}\ (1/m) & S_{yy}\ (1/m + 1/n) \end{bmatrix} = \text{sample covariance matrix of } [P]$$

$$[A] = \begin{bmatrix} a_y & a_x & a_d \\ \\ c_y & c_x & c_d \end{bmatrix} = -[H]'\ [G]^{-1} = \text{matrix of SPR coefficients}$$

The general SPR formulae, that are always the same (after the various matrices have been appropriately defined, can now be written as

(1) $[\hat{\mu}] = [Q] + [A]'\ [P] = [Q] - [H]'\ [G]^{-1}\ [P] =$ SPR estimator of $[\mu]$

and

(2) $[S_{\hat{\mu}\hat{\mu}}] = [K] + [A]'\ [H] = [K] - [H]'\ [G]^{-1}\ [H] =$ estimator of the covariance matrix of $[\hat{\mu}]$

Note that the SPR estimator consists of two parts. There is first a vector [Q] whose expected value is the parameter of interest, in this case, the vector $[\mu]$, and a second part, the adjustment [A]' [P] made to [Q] for the difference between the means of the matched and unmatched subsamples. Although the expected value of [P] is zero, the expected value of the adjustment [A]' [P] is not zero; this is because [A] and [P] are not statistically independent. Consequently, the SPR estimators are basically biased. To apply the matrix approach requires (i) the identification of two, appropriately defined vectors [Q] and [P] and (ii) the evaluation of their covariance matrices, something which is not as difficult as it may seem at first sight. Somebody with some knowledge of mathematical statistics can be easily trained to learn how to evaluate the covariance matrices [K], [H] and [P] and for today's computers, it is a relatively simple matter to do the required matrix calculations. Finally, note that one advantage of the matrix approach is that the covariance matrix $[S_{\hat{\mu}\hat{\mu}}]$ contains, in addition to the variances of the SPR estimators, all of their covariance terms, terms that can be found useful, sometimes.

It may be interesting here to show how to apply the SPR approach to numerical data. Using the sample data and the basic summary statistics as given in Cunia (1989 a) we can write

$$[Q] = \begin{bmatrix} 1.96818182 \\ 18.47222222 \\ -1.50404040 \end{bmatrix}, \quad [P] = \begin{bmatrix} 3.97710027 \\ \\ .86574280 \end{bmatrix}$$

$$[K] = \begin{bmatrix} 4.803988359 & 0 & 4.803988389 \\ 0 & 5.565887533 & -5.565887533 \\ 4.803988359 & -5.565887533 & 10.36987589 \end{bmatrix}$$

$$[H] = \begin{bmatrix} 0 & 5.565887533 & -5.565887533 \\ \\ 4.803988359 & 0 & 4.803988359 \end{bmatrix}$$

$$[G] = \begin{bmatrix} 8.009447914 & 2.492292980 \\ 2.492292980 & 7.381738210 \end{bmatrix}$$

As

$$[G]^{-1} = \begin{bmatrix} .1395093929 & -.04710249410 \\ -.04710249410 & .1513726420 \end{bmatrix}$$

we can write now

$$[A] = -[G]^{-1}[H] = \begin{bmatrix} .2262798333 & -.7764935907 & 1.002773424 \\ -.727192410 & .2621671847 & -.9893595947 \end{bmatrix}$$

and thus

$$[\hat{\mu}] = [Q]+[A]'[H] = \begin{bmatrix} 17.23855781 \\ 15.61099870 \\ 1.627559109 \end{bmatrix}$$

$$[S_{\hat{\mu}\hat{\mu}}] = [K] + [A]'[H] = \begin{bmatrix} 1.3105645 & 1.2594481 & .051116384 \\ 1.2594481 & 1.2440115 & .015436566 \\ .051116384 & .015436566 & .035679810 \end{bmatrix}$$

SPR ESTIMATORS WITH AUXILIARY VARIABLES

Recall the interpretation given by Cunia (1989a) to the SPR estimator y as a weighted mean of (i) $\bar{y}_d$, a double sampling estimator that used the information in the (u + m) units and (ii) $\bar{y}_n$ a simple random sampling estimator that uses the information in the remaining n new units of the second occation. It is known that, instead of the usual linear regression of y on x used in the definition of $\bar{y}_d$, one can use the multiple linear regression of y on x_2, x_2, ... , x_p, a set of p independent (auxiliary) variables x and, thus, define a multiple double sampling estimator. Futhermore, it is customary to measure the sample units for a host of variables, in addition to the variables x and y of specific interest, at a given application. All these variables are interrelated. As the SPR estimators implicitly use the linear regression coefficients $b_{y \cdot x}$ and $b_{x \cdot y}$, it seems reasonable to define SPR estimators that would implicitly use multiple linear regression coefficients.

More specifically, let us assume that (i) the (u + m) units are measured (on the first occasion) for the vector $[x]' = [x_1, x_2, \ldots, x_p]$, (ii) the (m + n) units are measured (on the second occasion) for the vector $[y]' = [y_1\ y_2, \ldots, y_q]$ and (iii) as before, the (u + m + n) units are all selected by simple random sampling. Then, let us define the SPR estimators with auxiliary variables of μ_{y1}, μ_{x1} and μ_d, which now will denote the means of y1, x1 and their difference d = (y1 - x1).

We start with the definition of the sample means

$$[\bar{x}_u] = \begin{bmatrix} \bar{x}_{1u} \\ \vdots \\ \bar{x}_{pu} \end{bmatrix}, \quad [\bar{x}_m] = \begin{bmatrix} \bar{x}_{1m} \\ \vdots \\ \bar{x}_{pm} \end{bmatrix}, \quad [\bar{y}_m] = \begin{bmatrix} \bar{y}_{1m} \\ \vdots \\ \bar{y}_{qm} \end{bmatrix}, \quad [\bar{y}_n] = \begin{bmatrix} \bar{y}_{1m} \\ \vdots \\ \bar{y}_{qn} \end{bmatrix}$$

and that of the covariances matrices (evaluated, as before, from the data of the m matched plots alone)

$$[S_{xx}] = \begin{bmatrix} S_{x1x1} & \cdots\cdots & S_{x1xp} \\ \vdots & & \vdots \\ S_{xpx1} & \cdots\cdots & S_{xpxp} \end{bmatrix}, \quad [S_{xy}] = \begin{bmatrix} S_{x1y1} & \cdots\cdots & S_{x1yq} \\ \vdots & & \vdots \\ S_{xpy1} & \cdots\cdots & S_{xpyq} \end{bmatrix}$$

and a similar expression for $[S_{yy}]$. All three SPR estimators have the same formulae

$$\hat{\mu} = [Q] - [H]'[G]^{-1}[P] \quad \text{and} \quad S_{\hat{\mu}\hat{\mu}} = [K] - [H]'[G]^{-1}[H]$$

where

(1) for the estimator $\hat{\mu} = \hat{y}_1$ of μ_{y1} we have

$$[Q_y] = [\bar{y}_{1n}]$$

$$[P_y]' = [(\bar{x}_{iu} - \bar{x}_{im}) \,\ldots\ldots\, (\bar{x}_{pu} - \bar{x}_{pm}) \;\; (\bar{y}_{1n} - \bar{y}_i)]$$

$$[K_y] = [S_{y1y1}/n]$$

$$[H_y]' = [[0]' \;\; S_{y1y1}/n]$$

where [0]' is a vector of p zero's and

$$[G_y] = \begin{bmatrix} [S_{xx}]\,(1/u + 1/m) & [S_{xy1}]\,(1/m) \\ \\ [S_{xy1}]'\,(1/m) & S_{y1y1}\,(1/m + 1/n) \end{bmatrix}$$

(2) for the estimator $\hat{\mu} = \hat{x}_1$ of μ_{x1} we have

$$[Q_x] = [\bar{x}_{iu}]$$

$$[P_x]' = [(\bar{x}_{1u} - \bar{x}_{1m}) \;\; (\bar{y}_{1n} - \bar{y}_{1m}) \,\ldots\ldots\ldots\, (\bar{y}_{qn} - \bar{y}_{qm})]$$

$$[K_x] = [S_{x1x1}/u]$$

$$[H_x] = [S_{x1x1}/n \;\;\; [0]']$$

where [0] is a vector of q zero's, and

$$[G_x] = \begin{bmatrix} S_{x1x1}\,(1/u + 1/m) & [S_{x1y}]](1/m) \\ \\ [S_{x1y}]'(1/m) & [S_{yy}](1/m + 1/n) \end{bmatrix}$$

and

(3) for the estimator $\hat{\mu} = \hat{d}$ of $\mu_d = (\mu_{y1} - \mu_{x1})$ we have

$$[Q] = [\bar{y}_{1n} - \bar{x}_{1u}] = [Q_y] - [Q_x]$$

$$[P]' = [\bar{x}_{1u} - \bar{x}_{1m}) \,\ldots\ldots\ldots\, (\bar{x}_{pu} - \bar{x}_{pm}) \;\; (\bar{y}_{1n} - \bar{y}_{1m}) \,\ldots\ldots\ldots\, (\bar{y}_{1n} - \bar{y}_{1m})]$$

$$[K] = [S_{x1x1}/u + S_{y1y1}/n] = [K_y] + [K_x]$$

$$[H] = [-[S_{xx1}]'/u \;\; [S_{yy1}]'/n]$$

and

$$[G] = \begin{bmatrix} [S_{xx}](1/u + 1/m) & [S_{xy}](1/m) \\ \\ [S_{xy}]'(1/m) & [S_{yy}](1/m + 1/n) \end{bmatrix}$$

It may be interesting to make a few remarks here. First, the subscripts of the variables x and y are arbitrarily selected and therefore, the estimators above can be defined for any variable x_i, variable y_j and difference $(y_j - x_i)$. Second, it can be shown that the SPR estimator of μ_{y1} that uses auxiliary variables, is the weighted mean of two other estimators: $\bar{y}_d$, the double sampling with multiple regression estimator using the data of the first (u + m) units and $\bar{y}_n$, the simple random sampling estimator defined on the last n units. Third, the problem of selecting the auxiliary variables can be solved by using multiple regression techniques; one can use only those variables x that are significant in the multiple linear regression of y_1 on $x_1, x_2, \ldots, x_p$ and only those variables y that are significant in the regression of x1 on $y_1, y_2, \ldots, y_q$. Finally, special rules can be easily derived for the case where one of the subsamples of unmatched (u and n) or matched (m) units is empty. These rules, however, are not shown here.

An important feature of the above SPR estimators is that they are not additive; the estimator $\hat{d}$ is not equal to $(\hat{y}_1 - \hat{x}_1)$. This is due to the fact that $[P_y]$, $[P_x]$ and [P] are differently defined. To make them additive, one can define a second set of SPR estimators $\hat{y}_1{}^*$ and $\hat{x}_1{}^*$ that use vectors $[P_y] = [P_x] = [P]$, estimators whose values are very close to those obtained above. Let us define these new estimators by working simultaneously with all three. For convenience we shall drop the superscript *.

Let the parameter to estimate be the vector

$$[\mu]' = [\mu_{y1} \;\; \mu_{x1} \;\; \mu_d]$$

where

$$\mu_d = \mu_{y1} - \mu_{x1}.$$

Then

$$[\hat{\mu}] = \begin{bmatrix} \hat{y}_1 \\ \hat{x}_1 \\ \hat{d} \end{bmatrix} = [Q] + [A]'[P] = [Q] - [H]'[G]^{-1}[P] = \text{SPR estimator of } [\mu]$$

and

$$[S_{\hat{\mu}\hat{\mu}}] = [K] + [A]'[H] = [K] - [H]'[G]^{-1}[H] = \text{covariance matrix of } [\mu]$$

where

$$[Q] = \begin{bmatrix} \bar{y}_{1n} \\ \bar{x}_{1u} \\ (\bar{y}_{1n} - \bar{x}_{1u}) \end{bmatrix} \quad , \quad [P] = \begin{bmatrix} [\bar{x}_u] - [\bar{x}_m] \\ \\ [\bar{y}_n] - [\bar{y}_m] \end{bmatrix}$$

$$[K] = \begin{bmatrix} S_{y1y1}/n & 0 & S_{y1y1}/n \\ 0 & S_{x1x1}/u & -S_{x1x1}/u \\ S_{y1y1}/n & -S_{x1x1}/u & (S_{y1y1}/n + S_{x1x1}/u) \end{bmatrix}$$

$$[H] = \begin{bmatrix} [0] & [S_{xx1}](1/u) & -[S_{xx1}](1/u) \\ \\ [S_{yy1}]'(1/n) & [0] & [S_{yy1}](1/n) \end{bmatrix}$$

and

$$[G] = \begin{bmatrix} [S_{xx}](1/u + 1/m) & [S_{xy}](1/m) \\ \\ [S_{xy}]'(1/m) & [S_{yy}](1/m + 1/n) \end{bmatrix}$$

We cannot give here an illustrative numerical example. The interested reader may find one in Cunia (1985).

MULTIVARIATE SPR ESTIMATORS

Sometimes one needs to estimate several parameters of interest, and in their definition of SPR estimators, it is possible to use the same vector [P] of differences between the mean values of [x] and [y] of the various subsamples. Then, it may be more convenient to estimate them all simultaneously, and obtain, estimates of covariances, in addition to those of their variances. Note that we have already done this when we have simultaneously derived the estimators of μ_y, μ_x and $\mu_d = (\mu_y - \mu_x)$. Now, however, we shall present the more general approach by referring to a specific case that will be needed in companion papers by Cunia (1989 b,c) containing also an illustrative numerical example.

Consider the variables

$$[x] = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \text{number of trees per acre} \\ \text{(sum of tree diameters) per acre} \\ \text{(sum of squared diameters) per acre} \end{bmatrix}$$

as measured on a set of (u + m) sample plots at the first measurement time and a similar vector [y] as measured on a set of (m + n) sample plots measured at the second measurement time, where the m plots are common to both occasions. As usual we define the vectors of sample means

$[\bar{x}_u]$, $[\bar{x}_m]$, $[\bar{y}_m]$, $[\bar{y}_n]$

and the sample covariance matrices (estimated from the m matched plots)

$[S_{xx}]$, $[S_{xy}]$, $[S_{yy}]$

Define also,

$$[Q_x] = [\bar{x}_u] = \begin{bmatrix} \bar{x}_{1u} \\ \bar{x}_{2u} \\ \bar{x}_{3u} \end{bmatrix} \quad \text{and} \quad [P] = \begin{bmatrix} [\bar{x}_u] - [\bar{x}_m] \\ \\ [\bar{y}_n] - [\bar{y}_m] \end{bmatrix}$$

Note that one can write [P]' as

$$[P]' = \left[(\bar{x}_{1u}-\bar{x}_{1m}) \; (\bar{x}_{2u}-\bar{x}_{2m}) \; (\bar{x}_{3u}-\bar{x}_{3m}) \; (\bar{y}_{1n}-\bar{y}_{1m}) \; (\bar{y}_{2n}-\bar{y}_{2m}) \; (\bar{y}_{3n}-\bar{y}_{3m})\right]$$

As the covariance matrices of $[Q_x]$ and [P] are

$$[K_x] = [S_{xx}]/u \quad , \quad [H] = \begin{bmatrix} [S_{xx}]/u & [S_{xy}]/m \\ \\ [S_{xy}]'/m & [S_{yy}]/n \end{bmatrix}$$

and

$$[G] = \begin{bmatrix} [S_{xx}]\,(1/u + 1/m) & [S_{xy}]\,(1/m) \\ [S_{xy}]'\,(1/m) & [S_{yy}]\,(1/m + 1/n) \end{bmatrix}$$

the SPR estimator of μ_{x1} , μ_{x2} and μ_{x3} and its variance is given respectively by

$$[x] = [Q_x] - [H_x]'\,[G]^{-1}\,[P] \quad \text{and} \quad [S_{xx}] = [K_x] - [H_x]'\,[G]^{-1}\,[H_x]$$

Similar procedures can be used to estimate separately $[\mu_y]$ or $[\mu_d] = [\mu_y] - [\mu_x]$ or to estimate simultaneously all nine parameters included in the three vectors $[\mu_x]$, $[\mu_y]$ and $[\mu_d]$, each vector being of order 3. These formulae are not shown here explicity.

SPR ESTIMATORS ON MORE THAN TWO OCCASIONS

The procedure, terminology and notation used here can be applied to the problem of calculating SPR estimators where there are three or more occasions. The size of the vector [P] may increase tremendously, however, since one must take into account all differences between subsample means of each occasion and between occasions. Because of space limitations we shall not consider the general problem, already considered in a formal way (although in a different notation) by Cunia and Chevrou (1969). What we shall do is to briefly outline, as an example, the estimator of μ_z, the mean of the variable of interest at the third measurement time. We shall assume that we have all combinations of units measured on one, two or three occasions. The number of sample units will denoted by n and its subscripts will denote the occastion they were measured.

More specifically we have n_1, n_2, n_3, n_{12}, n_{13}, n_{23} and n_{123} sample units and if x, y and z represents the variable of interest measured on the first, second and third occasions respectively, we shall denote the various subsample averages as x_1, y_2, z_3, x_{12}, y_{12}, x_{13}, z_{13}, y_{23}, z_{23}, x_{123}, y_{123} and z_{123}. As we want the SPR estimator of μ_z, we write

$$\hat{z} = [Q_z] + [A_z]'\,[P] = [Q_z] - [H_z]'\,[G]^{-1}\,[P]$$

and

$$S_{\hat{z}\hat{z}} = [K_z] + [A_z]'\,[H_z] = [K_z] - [H_z]'\,[G]^{-1}\,[H_z]$$

where

$$[Q_z] = [\bar{z}_3]$$

$$[P]' = [(x_1 - x_{123})\ (x_1 - x_{12})\ (x_1 - x_{13})\ (y_2 - y_{123})\ (y_2 - y_{12})\ (y_2 - y_{23})\ (z_3 - z_{123})\ (z_3 - z_{13})\ (z_3 - z_{23})]$$

$$[K_z] = [S_{zz}]/n_3$$

$$[H_z]' = [\,0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ S_{zz}/n_3\ \ S_{zz}/n_3\ \ S_{zz}/n_3]$$

and

$$[G] = \begin{bmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{bmatrix}$$

with

$$[G_{11}] = (S_{xx}) \begin{bmatrix} (1/n_1 + 1/n_{123}) & (1/n_1) & (1/n_1) \\ (1/n_1) & (1/n_1 + 1/n_{12}) & (1/n_1) \\ (1/n_1) & (1/n_1) & (1/n_1 + 1/n_{13}) \end{bmatrix}$$

similar expressions for $[G_{22}]$ and $[G_{33}]$, and

$$[G_{12}] = (S_{xy}) \begin{bmatrix} 1/n_{123} & 0 & 0 \\ 0 & 1/n_{12} & 0 \\ 0 & 1/n_{13} & 0 \end{bmatrix}$$

with similar expressions for the remaining submatrices.

REFERENCES

Cunia, T., 1985: More basic designs for survey sampling - Forest Biometry Monograph No. 4, SUNY College of Environmental Science and Foresty, Syracuse N.Y. 13210.

Cunia, T., 1989a: Successive forest inventories: basic SPR approach for two measurements. Proceedings of the State of the Art Methodology of Forest Inventory Symposium, July 31 - August 6 1989, SUNY College of Environmental Science and Forestry, Syracuse N.Y. 13210.

Cunia, T., 1989b: Successive forest inventories: error of biomass regressions - Part I: CFI wth permanent plots only. Proceedings of the State of the Art Methodology of Forest Inventory Symposium, July 31-August 6 1989, SUNY College of Environmental Science and Forestry, Syracuse N.Y. 13210.

Cunia, T., 1989c: Successive forest inventories: error of biomass regresssions - Part 2: CFI with partial replacement of plots. Proceedings of the State of the Art Methodology of Forest Inventory Symposium, July 31-August 6 1989, SUNY College of Environmental Science and Forestry, Syracuse N.Y. 13210.

Cunia, T. and Chevrou, R.B., 1969: Sampling with partial replacement on three or more occasions. Forest Science, Vol. 15, No. 2, pp. 205-224.

Successive Forest Inventories: Error of Biomass Regressions
Part I: CFI with Permanent Plots Only

Tiberius Cunia
Professor of Statistics and Operations Research
State University of New York College of Environmental Science and Forestry
Syracuse NY 13210

ABSTRACT

An approach is shown for the Continuous Forest Inventory (CFI) systems, to combine the error of the sample plots with that of the biomass (volume) regression functions. This approach requires that (i) the biomass regression functions be linear in form, with the covariance matrix of its regression coefficients available and (ii) the sample plots be selected by simple random sampling. Illustrative numerical examples show that the error of the biomass regressions compared to that of the sample plots is (i) relatively very small for the estimate of growth and change between various measurement values (ii) relatively important for the estimates of the current mean biomass/acre values and (iii) extremely large for the estimate of ingrowth. This seems to suggest that if one wants to have a realistic estimate of the error, he must take into account the error of the biomass regressions.

INTRODUCTION

The error of biomass estimates in a system of successive forest inventories has, like most inventory systems, two main components; one component due to the main sample of plots, permanent or temporary where the trees are measured for many attributes other than biomass and another component due to the relationship, biomass tables or regressions that translate the plot measurements into biomass averages per unit area or totals for the entire forest areas. Generally, the second error component is being ignored, either because it is assumed to be too small, or because the biomass regressions have no error statement (in appropriate form) attached to them or simply because one does not know how to take this error into account.

One procedure to take the second error component into account consists of the following main steps

Step 1 - The estimator of the parameter of interest μ is expressed as the statistic

$$w = [b]'[z] = b_1z_1 + b_2z_2 + \dots + b_mz_m$$

where

(i) $[b]$ = vector of estimates of the regression coefficients of the true biomass regression function which is of the form

$$E(y|x) = [\beta]'[x] = \beta_1x_1 + \beta_2x_2 + \dots + \beta_mx_m$$

= Expected value of the biomass y for given values of $[x]$

and

(ii) $[z]$ = vector of sample plot statistics
= estimator of the parameter $E([z]) = [\mu_z]$

such that, the parameter of interest μ is equal to the vector product $[\beta]'[\mu_z]$.

Step 2 - The error of the two vector statistics $[b]$ and $[z]$ is expressed as the sample covariance matrices $[S_{bb}]$ and $[S_{zz}]$, estimators of the true covariance matrices $[\sigma_{bb}]$ and $[\sigma_{zz}]$ respectively.

Step 3 - The error of w, expressed as the variance S_{ww}, the estimator of the true variance σ_{ww} of w is approximately given by

$$S_{ww} = [b]'[S_{zz}][b] + [z]'[S_{bb}][z] \text{ + terms of lower order of magnitude}$$

Remarks - (1) If $[z]$, $[b]$, $[S_{zz}]$ and $[S_{bb}]$ are all unbiased, then w and S_{ww} are also unbiased.

(2) First part of the expression S_{ww} is due to the error of the sample plots, the second part is due to the error of the biomass regression.

(3) It can be shown that in relative terms, (total percent standard error of w)^2 = (percent standard error of component 1)^2 + (percent standard error of component 2)^2.

The objectives of the present paper are to apply this approach to biomass estimates of successive forest inventory systems. We shall make the assumptions that the sample plots are selected by simple random sampling and the biomass regression functions are linear in form with (unbiased) estimates [b] and $[S_{bb}]$ statistically independent of the sample plot data. We shall also assume that [b] and $[S_{bb}]$ are given; we shall not be concerned with the problem of how to calculate them. The classical CFI system without SPR will be considered here by following very closely the methodology as described by Cunia (1987). The more general case of successive forest inventory with SPR, will be considered later in a companion paper, Cunia (1989). For simplicity we shall also consider only the case where the trees are of the same species.

Before considering the error of the main biomass estimators, let us define, in general, the variables of the cluster (sample plot or point) that we shall need. We shall assume given that the sample biomass regression is of the form

$$\hat{y} = b_1x_1 + b_2x_2 + \dots + b_mx_m$$

where (i) y is the tree biomass and $x_1 = 1, x_2, \dots, x_m$ are tree attributes such as diameter, height, (diameter)^2, etc. and (ii) [b] is an unbiased estimate of the $[\beta]$ and $[S_{bb}]$ an unbiased estimate of its covariance matrix.

The elements of the cluster are the trees and let Σ denote "summation over all trees in the cluster". Then, we define the following variables

$\hat{y}_{hk} = [b]'[x_{hk}] = b_1x_{hk1} + b_2x_{hk2} + \dots + b_mx_{hkm}$

= estimate of the biomass of the k-th tree in the h-th cluster

$\hat{y}_{hk}/a_{hk} = \hat{y}_{hk}$ expressed on a "per unit area" basis where a_{hk} represents the area of cluster corresponding to the tree hk

$\hat{v}_h = \Sigma\,(\hat{y}_{hk}/a_{hk})$ = estimate of the biomass "per unit area" of cluster h

If we define now the cluster variables as

$$s_{h1} = \Sigma\,(x_{hk1}/a_{hk})$$
$$s_{h2} = \Sigma\,(x_{hk2}/a_{hk})$$
$$\vdots$$
$$s_{hm} = \Sigma(x_{hkm}/a_{hk})$$

then, it can be shown that

$$\hat{v}_h = b_1s_{h1} + b_2s_{h2} + \dots + b_ms_{hm} = [b]'[s_h]$$

The summation Σ is taken over specific trees of the clusters. These trees are determined by the specific parameter μ one wishes to estimate. If the tree comes from a plot of fixed area "a", then, for all h and k

a_{hk} = a (acres or hectares).

But, if the tree comes from a point sample with BAF (basal area factor) of c, then

$a_{hk} = \Pi d^2_{hk}/4c$, if diameter d and BAF c are expressed in equivalent units

$= \Pi d^2_{hk}/40{,}000\ c$, if d is given in cm and c in m^2/ha of basal area

$= \Pi d^2_{hk}/576\ c$, if d is given in inches and c in sq. ft./acre of basal area

For example when the biomass regression is

$$\hat{y} = b_1 + b_2d + b_3d^2 = b_1x_1 + b_2x_2 + b_3x_3,$$

then, for a plot of fixed area "a"

$s_{h1} = (\Sigma\,(1)\,/a)$ = (number of trees per acre) in plot h

$s_{h2} = (\Sigma x_2/a) = \Sigma d/a$ = (sum of diameters per acre) in plot h

$s_{h3} = (\Sigma x_3/a) = \Sigma d^2/a$ = (sum of squared diameters/acre) in plot h

Note that s_{h3} is proportional to (basal area per acre) in plot h. In what follows, we shall use superscript 1 or 2 to denote measurement 1 or 2 of the sample tree. These superscripts will be dropped (for convenience) when there is no danger of confusion.

ESTIMATES OF THE ERROR OF THE MAIN PARAMETERS

(1) Estimator of μ_1 = mean biomass/unit area: measurement 1

We have already defined the cluster variables $s_1, s_2, ..., s_m$ as sums of the tree variables expressed on a per unit area basis and v_h as the biomass per unit area of the cluster h, h = 1, 2, ... , n. When $s_1, s_2, ..., s_m$ are calculated from the tree values of the first measurement, we have then

$$\bar{v} = \Sigma\hat{v}/n = (b_1\Sigma s_1 + b_2\Sigma s_2 + ... + b_m\Sigma s_m)/n = b_1\bar{s}_1 + b_2\bar{s}_2 + ... + b_m\bar{s}_m = [b]'[\bar{s}]$$

with the obvious definition of the average vector $[\bar{s}]$. Define $[z^1] = [\bar{s}]$ and $[b]'[z^1] = w_1$ = estimator of μ_1. As $S^1_{z_iz_j} = S_{\bar{s}_i\bar{s}_j} = S_{s_is_j}/n$, where $S_{s_is_j}$ is the sample covariance of the cluster variables s_i and s_j, we can easily define $[S^1_{zz}]$ = covariance matrix of $[z^1]$, and thus,

$w_1 = [b]'[z^1]$ = estimator of μ_1, the mean biomass per unit area of measurement 1, and

$S_{w1w1} = [b]'[S^1_{zz}][b] + [z^1]'[S_{bb}][z^1]$ = estimator of the variance of w_1.

(2) Estimator of μ_2 = mean biomass/unit area: measurement 2

Working with the second, rather than the first measurement values, we define $[z^2]$ and $[S^2_{zz}]$. Then,

$w_2 = [b]'[z]$ = estimator of μ_2, the mean biomass per unit area of measurement 2, and

$S_{w2w2} = [b]'\{S^2_{zz}][b] + [z^2]'[S_{bb}][z^2]$ = estimator of the variance of w_2

(3) Estimator of $\mu_g = (\mu_2-\mu_1)$ = mean change in biomass/unit area between measurement 1 and 2

The obvious estimator is

$$w_g = w_2-w_1 = [b]'[z^2] - [b]'[z^1] = [b]'[z^g]$$

where

$$[z^g] = [z^2] - [z^1] = [\bar{s}^2] - [\bar{s}^1] = [\bar{s}^g]$$

Note that the i-th element of $[z^g]$ is

$$z^g_i = \Sigma (s^2_{hi} - s^1_{hi})/n = \Sigma s^g_{hi}/n = \bar{s}^g_i$$

Each plot has m values s^g_{hi} that yield an average vector $[\bar{s}^g] = [z^g]$ and a covariance matrix $[S^g_{ss}]$ that defines $[S^g_{zz}] = [S^g_{ss}]/n$. This implies that

$w_g = [b]'[z^g]$ = estimator of μ_g = mean change

and

$S_{wgwg} = [b]'[S^g_{zz}][b] + [z^g]'[S_{bb}][z^g]$ = estimator of the variance of w_g

(4) Estimators of: μ_s = mean growth/unit area for the survivor trees only,
μ_m = mean mortality/unit area (biomass as recorded in measurement 1)
μ_i = mean ingrowth/unit area (biomas as recorded in measurement 2)

To define the estimator w_s of μ_s, we shall use the same formulae as in the definition of w_g, with the exception that in the calculation of the variables s we use only the survivor trees, those trees that are present and merchantable at both measurements 1 and 2. To define the estimator w_m of μ_m, we shall use the same formulae as in the definition of w_1, with the exception that in the calculation of the variables s we use only the trees that have died between the two measurements. Finally, by using only the ingrowth trees and the formulae of w_2, we define the estimator w_i of μ_i. The explicit formulae of w_s, w_m and w_i, and their error are not shown here.

The formulae above apply to the case where (i) the clusters are sample plots of fixed area or relascope sample points selected by simple random sampling and (ii) the biomass regression functions are linear in form and estimated separately and statistically independent of the sample cluster data. The formulae for the estimation of μ_t = mean volume per tree and those for the case where we separate the trees by classes (species) with different biomass regressions for each class are given by Cunia (1987). These are not considered here.

ILLUSTRATIVE NUMERICAL EXAMPLE

Consider a sample of 118 one-tenth acre plots measured on two occasions, in 1960 and 1964. If d is the tree diameter (in inches), its biomass y (in pounds of green weight) is estimated by the regression function

$$\hat{y} = b_1 + b_2 d + b_3 d^2 = b_1 x_1 + b_2 x_2 + b_3 x_3 = [b]'[x]$$

where

$$[b] = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 5.1818118 \\ -25.653078 \\ 12.988357 \end{bmatrix}, \text{ and } [S_{bb}] = \begin{bmatrix} 8715.8855 & -2222.4882 & 128.69992 \\ -2222.4882 & 581.99570 & -34.776995 \\ 128.69992 & -34.776995 & 2.1744582 \end{bmatrix}$$

The calculation of these statistics from the data of the 353 sample trees is shown in Cunia (1986).

Using the estimated tree biomass values, we calculate then (by summation) the estimate of the biomass per acre of each given plot h, h = 1, 2, ..., 118 by the formula

$\hat{v}_h$ = sum of the estimates of the biomass of the n_h trees of plot h

$$= 10\,(\hat{y}_{h1} + \hat{y}_{h2} + \ldots + \hat{y}_{hn}) = 10\,\Sigma\,\hat{y}_{hk}$$

where 10 is the converting factor from the "per plot" to the "per acre" value. As

$\hat{y}_{hk} = b_1 + b_2 d_{hk} + b_3 d^2_{hk}$ = estimated biomass of the k-th tree in the h-th plot, we can write

$$\hat{v}_h = 10\,\Sigma\,(b_1 + b_2 d_{hk} + b_3 d^2 hk) = b_1\,(10 n_h) + b_2\,(10\,\Sigma\, d_{hk}) + b_3\,(10\,\Sigma\, d^2_{hk})$$
$$= b_1 s_{h1} + b_2 s_{h2} + b_3 s_{h3} = [b]'[s_h]$$

Using now the tree diameters of the 118 plots, we calculate the values s_1, s_2 and s_3 of the first (1960) and the second (1964) measurements. These are given by Cunia (1986) in Table 2 (page 13) and Table 3 (page 31); these are not repeated here. Note that

s_1 = number of trees/acre of a given plot

s_2 = (sum of tree diameters/acre) of a given plot

s_3 = (sum of squared diameters/acre) of a given plot = proportional to basal area/acre

Using now the s_1, s_2 and s_3 values of the individual plots, we calculate

$\bar{v}$ = estimated average biomass/acre = $\Sigma\,\hat{v}_h/118$

$$= \Sigma\,(b_1 s_{h1} + b_2 s_{h2} + b_3 s_{h3})/118 = b_1 \bar{s}_1 + b_2 \bar{s}_2 + b_3 \bar{s}_3 = [b]'[\bar{s}]$$

The covariance of s_i and s_j is estimated by

$$S_{s_i s_j} = \Sigma(s_i - \bar{s}_i)(s_j - \bar{s}_j)/117$$

and that of $\bar{s}_i$ and $\bar{s}_j$ is estimated by

$$S_{\bar{s}_i \bar{s}_j} = S_{s_i s_j}/118$$

If we use the diameters of the first (1960) measurement, we can write

$$[\bar{s}] = [z^1] = \begin{bmatrix} 236.35593 \\ 1439.3161 \\ 10898.164 \end{bmatrix} \text{ and } [S^1 zz] = \begin{bmatrix} 180.27909 & 936.03701 & 5104.6720 \\ 936.03701 & 5584.8417 & 40038.488 \\ 5104.6720 & 40038.488 & 411898/95 \end{bmatrix}$$

The diameters of the second (1964) measurement yield

$$[\bar{s}] = [z^2] = \begin{bmatrix} 266.01695 \\ 638.6805 \\ 12320.176 \end{bmatrix} \text{ and } [S^2 zz] = \begin{bmatrix} 203.72504 & 1047.0999 & 5389.2817 \\ 1047.0999 & 6087.4263 & 40700.319 \\ 5389.2817 & 40700.319 & 398306.82 \end{bmatrix}$$

Finally, applying the standard formulae we find

$w_1 = [b]'[z^1] = 105{,}851$ pounds/acre = estimate of the mean biomass/acre in 1960

$w_2 = [b]'[z^2] = 119{,}360$ pounds/acre = estimate of the mean biomass/acre in 1964

$$S_{w1w1} = [b]'[S^1_{zz}][b] + [z^1]'[S_{bb}][z^1] = 46{,}923{,}716 + 10{,}710{,}584 = 57{,}634{,}300$$

= estimate of the variance of w_1, and

$$S_{w2w2} = [b]'[S_{zz}][b] + [z^2]'\,[S_{bb}][z^2] = 44{,}529{,}880 + 11{,}393{,}835 = 55{,}923{,}716$$

= estimate of the variance of w_2

Note that the second error component (due to the biomass regressions) represents about 20 percent of the total error, a relatively small percentage. However, it should be realized that this small percentage is obtained with 353 sample trees(for biomass regression construction) and 118 sample plots. If hypothetically, we have 88 trees (one quarter of the 353 number) and 472 plots (four times the number 118) then, the percent of the error component due to biomass regression would represent 80 percent of the total error.

To estimate the average net change between the first and the second measurement we use the obvious statistic

$$w_g = w_2 - w_1 = [b]'[z^2] - [b]'[z^1] = [b]'[z^g]$$

where

$$[z^g] = \begin{bmatrix} \bar{s}^2{}_1 - \bar{s}^1{}_1 \\ \bar{s}^2{}_2 - \bar{s}^1{}_2 \\ \bar{s}^2{}_3 - \bar{s}^1{}_3 \end{bmatrix} = \begin{array}{l} \text{average difference in number of trees/acre} \\ \text{average difference in (sum of diameters)/acre} \\ \text{average difference in (sum of sq. diam.)/acre} \end{array}$$

By working directly with the plot differences (expressed on a "per acre" basis), one can easily find the covariance terms needed to define

$[S^g_{zz}]$ = covariance matrix of $[z^g]$.

Then, Cunia (1986) shows that,

$$[z^g] = \begin{bmatrix} 29.661017 \\ 199.36441 \\ 1422.0114 \end{bmatrix} \text{ and } [S^g_{zz}] = \begin{bmatrix} 11.3274 & 54.4183 & 305.339 \\ 54.4183 & 301.110 & 2126.62 \\ 305.339 & 2126.62 & 21174.2 \end{bmatrix}$$

$$w_g = [b]'[z^g] = 13{,}509 = w_2 - w_1$$

and

$$S_{wgwg} = 2{,}379{,}986 + 50{,}666 = 2{,}430{,}655$$

Note that the second error component (due to the biomass regression) represents only about 2 percent of the total error, a negligible amount. This is due to the fact that (1) the values $[z^g]$ of the second error term, that is, $[z^g]'[S_{bb}][z^g]$ are differences between plot measurements four years apart, and as such are very small values, and (2) the values of the covariance terms of the first error term, that is, $[b]'[S_{zz}][b]$ are relatively large, again, because the differences z^g between plot values measured four years apart are very small.

Cunia (1986) gives the values s_1, s_2 and s_3 calculated (i) from the data of mortality trees only (first measurement) in Table 4 (page 32) and (ii) from the ingrowth trees only (second measurement) in Table 5 (page 33). The variable s_1, s_2 and s_3 of the growth (or changes) occurring on the survivor trees only (trees alive on both occasions) have been calculated from the data of Tables 3, 4 and 5 mentioned above. Then, it can be shown that

$w_s = 14718.385$ pounds/acre of net change in the biomass/acre of the survivor trees

$w_m = 4874$ pounds/acre of mortality between 1960 and 1964

$w_i = 3665$ pounds/acre of ingrowth between 1960 and 1964 with variances

$S_{wsws} = 439{,}839 + 608{,}965 = 1{,}048{,}804$

$S_{wmwm} = 1{,}996{,}418 + 24{,}567 = 2{,}020{,}984$

and

$S_{wiwi} = 98{,}330 + 767{,}729 = 866{,}052$

As the reader can verify, the second error component represents, in terms of the total error (i) some 60 percent for the growth on survivor trees (ii) some 1 percent for mortality and (iii) some 90 percent for ingrowth. It may be difficult to explain why the relative size of the second error component is so small for mortality; probably because the difference between plot mortality is extremely large, and this makes the error of biomass regression relatively small. On the other hand, the second error component of ingrowth is relatively large. This may be due to the fact that the ingrowth trees are at one side of the biomass regression function, where the error is the largest; especially true in relative terms, since the biomass of ingrowth is very small.

REFERENCES

Cunia, T. (1985): On the error of biomass estimates in forest inventory: Part I: Its major components. Faculty of Forestry Miscellaneous Publication No. 8 SUNY College of Environmental Science and Forestry - Syracuse NY.

Cunia, T. (1986): On the error of biomass estimates in forest inventories: Part 2: The error component from sample plots. Faculty of Forestry Miscellaneous Publication No. 9, SUNY College of Environmental Science and Forestry - Syracuse NY.

Cunia, T. (1987): On the Error of Continuous Forest Inventory Estimates. Canadian Journal of Forest Research, Vol. 17, No. 5, pp. 436-441.

Cunia, T. (1989): Successive forest inventories: error of biomass regressions - Part 2: CFI with partial replacement of plots. Proceedings of the "State of the Art Methodology of Forest Inventory" Symposium, July 31 - August 6, 1989, SUNY College of Environmental Science and Forestry - Syracuse NY.

Successive Forest Inventories: Error of Biomass Regressions
Part 2: CFI with Partial Replacement of Plots

Tiberius Cunia
Professor of Statistics and Operations Research
State University of New York College of Environmental Science and Forestry
Syracuse NY 13210

ABSTRACT

An approach is shown, for the Continuous Forest Inventory (CFI) systems that use Sampling with Partial Replacement (SPR) of plots, to combine the error of sample plots with the error of biomass (volume) regression functions. Applied to an illustrative numerical example, the approach shows that in many cases the error component due to the biomass regression is important enough, and cannot be ignored, if a valid estimate of the total error of the inventory estimates is desired.

INTRODUCTION

In a previous paper, Cunia (1989 c) described an approach to combine the error of the main sample of plots from a given inventory system, with the error of the biomass (volume) regression functions used to estimate the sample tree biomass. This approach was then applied to a Continuous Forest Inventory (CFI) design where the same sample plots are being measured and remeasured at regular time intervals. This approach requires in addition that (i) the biomass regression functions be linear, with an estimate of the covariance matrix of its coefficients available and (ii) the sample plots be selected by simple random sampling. In this paper we shall extend the methodology to the case where some of the plots at one measurement are being replaced by another set at remeasurement time, that is, the continuous inventory sampling designs that use sampling with partial replacement or SPR. We shall follow very closely the methodology as described by Cunia (1987).

For the general SPR approach the reader is referred to Cunia (1989 a, b) and for the approach to combine the error of the sample plots with that of the biomass regression function to Cunia (1989 c). The reader must be familiar with the terminology, notation and methodology of these papers before embarking in the study of the present one.

ESTIMATES OF THE ERROR OF THE MAIN PARAMETERS

Let us consider now the basic case of a CFI system with SPR, where the forest area is measured on two occasions only. The sample plots are measured on the first occasion for the vector $[x]' = [x_1 \ \ x_2 \dots x_p]$, where x_1 = plot biomass per acre, and on the second occasion for the vector $[y]' = [y_1 \ \ y_2 \dots y_q]$, where y_1 = plot biomass/acre. The objective of the sampling is to estimate μ_{x1} = mean biomass/acre at first occasion, μ_{y1} = mean biomass/acre at second occasion, and $\mu_g = (\mu_{y1} - \mu_{x1})$ = mean change in biomass/acre from the first to the second occasion. The variables $x_2, x_3, \dots, x_p$ and $y_2, y_3, \dots, y_q$ are known as auxiliary variables.

The basic sample data (that will be used to estimate μ_{y1}, μ_{x1} and μ_g) consists of (1) m permanent sample plots measured on both occasions; these data provided estimates (i) $[\bar{x}_m]$ and $[\bar{y}_m]$ of $[\mu_x]$ and $[\mu_y]$ and (ii) $[S_{xx}]$, $[S_{yy}]$ and $[S_{xy}]$ of the covariance matrices of [x] and [y] respectively, and (2) u plots measured for [x] but not for [y] (that is, plots measured only on the first occasion) and n plots measured for [y] but not for [x] (that is plots measured only on the second occasion); these data provide estimates $[\bar{x}_u]$ and $[\bar{y}_n]$ respectively of the $[\mu_x]$ and $[\mu_y]$.

Using matrix notation, we define now the usual SPR estimators, that we shall now call the "classical" SPR estimators of μ_{x1}, μ_{y1} and μ_g as

$$[w] = \begin{bmatrix} \hat{x}_1 \\ \hat{y}_1 \\ \hat{g} \end{bmatrix} = [Q] + [A]'[P]$$

where

$$[Q] = \begin{bmatrix} \bar{x}_{1u} \\ \bar{y}_{1n} \\ \bar{y}_{1n}-\bar{x}_{1u} \end{bmatrix} = \text{vector of averages of unmatched (u) and new (n) plots}$$

$$[P] = \begin{bmatrix} [\bar{x}_u] - [\bar{x}_m] \\ \\ [\bar{y}_n] - [\bar{y}_m] \end{bmatrix} = \begin{array}{l}\text{vector of differences between averages} \\ \text{of matched and unmatched (and new) plots}\end{array}$$

and

$$[A] = -[G]^{-1}[H] = \text{the } (p + q) \text{ by 3 matrix of SPR coefficients}$$

where

$$[G] = \begin{bmatrix} (u + m)[S_{xx}]/um & [S_{xy}]/m \\ \\ [S_{xy}]'/m & (m + n)[S_{yy}]/mn \end{bmatrix} = \text{covariance matrix of } [P]$$

$$[G]^{-1} = \text{inverse of } [G]$$

and

$$[H] = \begin{bmatrix} S_{x1x1}/u & 0 & -S_{x1x1}/u \\ \vdots & \vdots & \vdots \\ S_{x1xp}/u & 0 & -S_{x1xp}/u \\ 0 & Sy1y1/n & Sy1y1/n \\ \vdots & \vdots & \vdots \\ 0 & S_{y1yq}/n & S_{y1yq}/r \end{bmatrix} = \text{covariance matrix of } [P] \text{ and } [Q]$$

If we define in addition

$$[K] = \begin{bmatrix} S_{x1x1}/u & 0 & -S_{x1x1}/u \\ 0 & S_{y1y1}/n & S_{y1y1}/n \\ -S_{x1x1}/u & S_{y1y1}/n & (S_{x1x1}/u + S_{y1y1}/n) \end{bmatrix} = \text{covariance matrix of } [Q]$$

then, the estimator of the covariance matrix of w is

$$[S_{ww}] = [K] + [A]'[H] = [K] - [H]'[G]^{-1}[H] = \begin{bmatrix} S_{\hat{x}\hat{x}} & S_{\hat{x}\hat{y}} & S_{\hat{x}\hat{g}} \\ S_{\hat{x}\hat{y}} & S_{\hat{y}\hat{y}} & S_{\hat{y}\hat{g}} \\ S_{\hat{x}\hat{g}} & S_{\hat{y}\hat{g}} & S_{\hat{g}\hat{g}} \end{bmatrix}$$

Note that different SPR estimators are obtained whenever different sets of auxiliary variables are used. They all ignore the error of biomass regressions, when the covariance matrix $[S_{ww}]$ is calculated.

To take into account the error of the biomass tables, we need plot variables $s_1, s_2, \ldots$ of the type defined by Cunia (1989 c). If the biomass regression function of u = tree biomass on the tree variables $v_1, v_2, \ldots, v_r$ is of the form

$$u = b_1v_1 + b_2v_2 + \ldots + b_rv_r = [b]'[v]$$

with $v_1 = 1$, and we have the estimate $[S_{bb}]$ of the covariance matrix of [b], then the plot variables can be written as

$$s_i = \Sigma\, (v_{ik}/a_k)\ ,\quad 1 = 1, 2, \ldots, r\ ;\ \ k = 1, 2, \ldots, \text{number of trees/plot}$$

Each plot of the CFI with SPR system is measured then for the variables $s_1, s_2, \ldots, s_r$; (1) on the first occasion only (for the u unmatched plots); 2) on the second occasion only (for the n new plots); and (3) on both, first and second occasion (for the m matched plots). The plot data are now summarized as

(1) the vector of sample averages $[\bar{s}_u]$, $[\bar{s}^1{}_m]$, $[\bar{s}^2{}_m]$, $[\bar{s}_n]$ and (2) the covariance matrices of $[s^1]$ and $[s^2]$, calculated from the data of the m matched plots alone, $[S^{11}{}_{ss}]$, $[S^{12}{}_{ss}]$, $[S^{22}{}_{ss}]$.

To define the "new" SPR estimators, we start with

(1) the vectors of sample averages and differences

$$[Q_1] = [\bar{s}_u]\,,\ [Q_2] = [\bar{s}_n]\,,\ [Q_g] = [Q_2] - [Q_1] \text{ and } [P] = \begin{bmatrix} [\bar{s}_u] - [\bar{s}^1{}_m] \\ \\ [\bar{s}_n] - [\bar{s}^2{}_m] \end{bmatrix}$$

(2) the covariance matrices of the vectors above

$[K_1] = [S^{11}{}_{ss}]/u$ = estimates of the covariance matrix of $[Q_1]$

$[K_2] = [S^{22}{}_{ss}]/n$ = estimate of the covariance matrix of $[Q_2]$

$[K_g] = [K_1] + [K_2]$ = estimate of the covariance matrix of $[Q_g]$

$$[G] = \begin{bmatrix} (u+m)\,[S^{11}{}_{ss}]/um & [S^{12}{}_{ss}]/m \\ \\ [S^{12}{}_{ss}]'/m & (m+n)[S^{22}{}_{ss}]/mn \end{bmatrix} = \text{estimate of the covariance matrix of } [P]$$

$$[H_1] = \begin{bmatrix} [K_1] \\ \\ [0] \end{bmatrix} = \text{estimate of the covariance matrix of } [P] \text{ and } [Q_1]$$

$$[H_2] = \begin{bmatrix} [0] \\ \\ [K_2] \end{bmatrix} = \text{estimate of the covariance matrix of } [P] \text{ and } [Q_2\}$$

and finally

$$[H_g] = \begin{bmatrix} -[K_1] \\ \\ [K_2] \end{bmatrix} = \text{estimate of the covariance matrix of } [P] \text{ and } [Q_g]$$

We define now the SPR estimators of the vector of the expected values of $s_1, s_2, \ldots, s_r$, say $[\mu^1{}_z]$ and $[\mu^2{}_z]$, for the first and second measurements respectively, and their difference, as

$[z^1] = [Q_1] + [A_1]'[P]$ = estimate of $[\mu^1{}_z]$

$[z^2] = [Q_2] + [A_2]'\,[P]$ = estimate of $[\mu 2_z]$

$[z^g] = [Q_g] + [A_g]'[P]$ = estimate of $[\mu^1{}_z] - [\mu^2{}_z]$

where

$[A_1] = -[G]^{-1}[H_1]$, $[A_2] = -[G]^{-1}\,[H_2]$ and $[A_g] = -[G]^{-1}[H_g] = [A_2] - [A_1]$

are the matrices of SPR coefficients.

The covariance matrices of $[z^1]$, $[z^2]$ and $[z^g]$ are estimated respectively by

$[S^{11}{}_{zz}] = [K_1] + [A_1]'[H_1]$, $[S^{22}{}_{zz}] = [K_2] + [A_2]'[H_2]$, and $[S^{gg}{}_{zz}] = [K_g] + [A_g]'[H_g]$

Consequently, the SPR estimators of μ_1, μ_2 and μ_g are respectively given by

$w_1 = [b]'[z^1]$, $w_2 = [b]'[z^2]$ and $w_g = [b]'[z^g]$

with the usual variance formulae of the form

$w_i = [b]'[S^{ii}{}_{zz}][b] + [z^i]'[S_{bb}][z^i]$, where i = 1, 2 and g.

ILLUSTRATIVE NUMERICAL EXAMPLE; CFI WITH SPR

To the m = 118 permanent plots measured on two occasions of the Cunia (1989 c) example, we have added u = 117 temporary plots measured on the first (1960) occasion only and n = 92 temporary plots measured on the second (1964) occasion only. The biomass of a tree is still estimated by the same regression function

$$u = b_1 + b_2 d + b_3 d^2 = b_1 v_1 + b_2 v_2 + b_3 v_3 = [b]'[v]$$

where, u = tree biomass (do not confuse with u = number of plots above), d = tree diameter, and the numerical values of [b] and $[S_{bb}]$, reported previously by Cunia (1989 c) are not repeated here.

To calculate the classical and the new SPR estimates of μ_1 = mean biomass/acre of first measurement, μ_2 = mean biomass/acre of second measurement and $\mu_g = \mu_2 - \mu_1$, we need the following plot values: s_1 = number of trees/acre, s_2 = sum of diameters/acre and s_3 = sum of squared diameters/acre for both 1960 and 1964 measurements. The values of the individual plots are listed by Cunia (1986) in Table 1 for the u = 117 temporary plots (1960 values), Tables 2 and 3 for the m = 118 permanent plots (1960 and 1964 values) and finally Table 6 for n = 92 temporary plots (1964 values).

We have used the following classical and new SPR procedures. We only give the summary results calculated from the tables of individual plot values listed above.

Procedure 1 - The classical SPR estimates using no auxiliary variable other than the plot biomass/acre x (of 1960) and y (of 1964). The statistics are as follows

$$[Q] = \begin{bmatrix} \bar{x}_u \\ \bar{y}_n \\ \bar{y}_n - \bar{x}_u \end{bmatrix} = \begin{bmatrix} 106046.85 \\ 125267.64 \\ 19220.790 \end{bmatrix};\quad [P] = \begin{bmatrix} \bar{x}_u - \bar{x}_m \\ \\ \bar{y}_n - \bar{y}_m \end{bmatrix} = \begin{bmatrix} 195.73999 \\ \\ 5907.5513 \end{bmatrix}$$

$$[K] = \begin{bmatrix} S_{xx}/u & 0 & -S_{xx}/u \\ 0 & S_{yy}/n & S_{yy}/n \\ -S_{xx}/u & S_{yy}/n & (S_{xx}/u + S_{yy}/n) \end{bmatrix} = \begin{bmatrix} 47324773 & 0 & -57114411 \\ 0 & 57114411 & 57114411 \\ -47324773 & 57114411 & 104439185 \end{bmatrix}$$

$$[G] = \begin{bmatrix} (u+m)\, S_{xx}/um & S_{xy}/m \\ \\ S_{xy}/m & (m+n)\, S_{yy}/mn \end{bmatrix} = \begin{bmatrix} 94248489 & 44536804 \\ \\ 44536804 & 101644292 \end{bmatrix}$$

$$[H] = \begin{bmatrix} S_{xx}/u & 0 & -S_{xx}/u \\ \\ 0 & S_{yy}/n & S_{yy}/n \end{bmatrix} = \begin{bmatrix} 47324773 & 0 & -47324773 \\ \\ 0 & 57114411 & 57114411 \end{bmatrix}$$

$$[A] = -[G]^{-1}[H] = \begin{bmatrix} -.63324208 & .33485978 & .96810186 \\ \\ .27746347 & -.70862804 & -.98609152 \end{bmatrix}$$

Consequently, the classical SPR estimates are

$$[w] = \begin{bmatrix} w_1 \\ w_2 \\ w_g \end{bmatrix} = [Q] + [A]'[P] = \begin{bmatrix} 107562.03 \\ 121146.93 \\ 13584.90 \end{bmatrix}, \text{ and}$$

$$[S_{ww}] = [K] + [A]'[H] = \begin{bmatrix} 17356735 & 15847163 & -1509572 \\ 15847163 & 16641538 & 794375 \\ -1509572 & 794375 & 2303947 \end{bmatrix}$$

Procedure 2 - The new SPR estimates using the auxiliary variables s_1, s_2 and s_3 of the first and second measurements. The basic statistics of the (u+m+n) plots are:

$$[\bar{s}_u] = \begin{bmatrix} 214.78632 \\ 1338.3291 \\ 10722.382 \end{bmatrix} = \text{vector of 1960 averages calculated from the data of u = 117 plots}$$

$$[\bar{s}_n] = \begin{bmatrix} 228.69565 \\ 1502.8380 \\ 12521.600 \end{bmatrix} = \text{vector of the 1964 averages calculated from the data of } n = 92 \text{ plots}$$

$$[\bar{s}^1{}_m] = \begin{bmatrix} 236.35593 \\ 1439.3161 \\ 10898.164 \end{bmatrix} = \text{vector of the 1960 averages calculated from the data of } m = 118 \text{ plots}$$

$$[\bar{s}^2{}_m] = \begin{bmatrix} 266.01695 \\ 1638.6805 \\ 12320.176 \end{bmatrix} = \text{vector of the 1964 averages calculated from the data of } m = 118 \text{ plots}$$

$$[S^{11}{}_{ss}] = \begin{bmatrix} 21272.932 & 110452.37 & 602351.30 \\ 110452.37 & 659011.32 & 4724541.6 \\ 602351.30 & 4724541.6 & 48604076 \end{bmatrix} = \text{covariance matrix calculated from the data of the } m = 118 \text{ plots}$$

$$[S^{12}{}_{ss}] = \begin{bmatrix} 21987.926 & 118295.97 & 659678.73 \\ 109292.82 & 670898.34 & 4832348.2 \\ 542577.80 & 4443890.0 & 46552860 \end{bmatrix} = \text{covariance matrix of } s^1,\ s^2 \text{ and } s^3 \text{ of the first and second measurements}$$

$$[S^{22}{}_{ss}] = \begin{bmatrix} 24039.555 & 123557.78 & 635935.23 \\ 123557.78 & 718316.31 & 4802637.7 \\ 635935.23 & 4802637.7 & 47000205 \end{bmatrix} = \text{covariance matrix calculated from the data of the } m = 118 \text{ plots}$$

Using the basic statistics above we define the following SPR statistics

$$[Q_1] = [\bar{s}_u],\quad [Q_2] = [\bar{s}_n],\quad [Q_g] = [Q_2] - [Q_1],\quad [P] = \begin{bmatrix} [\bar{s}_u] - [\bar{s}^1{}_m] \\ [\bar{s}_n] - [\bar{s}^2{}_m] \end{bmatrix}$$

$$[K_1] = [S^{11}{}_{ss}]/u,\quad [K_2] = [S^{2}{}_{ss}]/n,\quad [K_g] = [K_1] + [K_2]$$

$$[G] = \begin{bmatrix} (u+m)[S^{11}{}_{ss}]/um & [S^{12}{}_{ss}]/m \\ [S^{12}{}_{ss}]'/m & (m+n)[S^{22}{}_{ss}]/mn \end{bmatrix}$$

$$[H_1] = \begin{bmatrix} [K_1] \\ [0] \end{bmatrix},\quad [H_2] = \begin{bmatrix} [0] \\ [K_2] \end{bmatrix},\quad [H_g] = \begin{bmatrix} -[K_1] \\ [K_2] \end{bmatrix}$$

$$[A_1] = -[G]^{-1}[H_1] = \begin{bmatrix} -.57582423 & .27162431 & 2.1942318 \\ -.016295496 & -.71115549 & -.60394450 \\ .00082660800 & .0038778696 & -.60368650 \\ -.052395398 & -1.41966765 & -6.1216962 \\ .075951873 & .58362944 & 1.2028896 \\ -.0031360528 & -.011748779 & .23558038 \end{bmatrix}$$

$$[A_2] = -[G]^{-1}[H_2] = \begin{bmatrix} .58503065 & 1.0689709 & 1.9045323 \\ -.051364277 & .14848762 & .068781280 \\ .0017007641 & .0047393886 & .30871816 \\ -.62752621 & .39234747 & 2.7713606 \\ -.021883489 & -.81539047 & -.73733567 \\ .0011022853 & .0052754490 & -.67282885 \end{bmatrix}$$

and $[A_g] =-[G]^{-1} [H_g] = [A_2] - [A_1]$

The above statistics lead directly to the SPR estimates of $[\mu_z]$, the vector of

$$\begin{bmatrix} z1 \\ z2 \\ z3 \end{bmatrix} = \begin{bmatrix} \text{mean number of trees/acre} \\ \text{mean sum of diameters/acre} \\ \text{mean sum of squared diameters/acre} \end{bmatrix}$$

of the first (1960) measurement time, say $[\mu^1{}_z]$, of the second (1964) measurement time, say $[\mu^2{}_z]$, and of their difference $[\mu^g{}_z] = [\mu^2{}_z] - [\mu^1{}_z]$. More specifically we have

$$[z^1] - [Q_1] + [A_1]' [P] = \begin{bmatrix} 219.71326 \\ 1374.9418 \\ 10954.679 \end{bmatrix} = \text{estimate of } [\mu^1{}_z]$$

$$[z^2] = [Q_2] + [A_2]' \{P] = \begin{bmatrix} 247.57976 \\ 1561.1367 \\ 12280.514 \end{bmatrix} = \text{estimate of } [\mu^2{}_z]$$

$$[z^g] = [z^2] - [z^1] = \begin{bmatrix} 27.86650 \\ 186.19492 \\ 1325.8344 \end{bmatrix} = \text{estimate of the mean change } [\mu^g{}_z]$$

and the estimates of their covariance matrices

$$[S^{11}{}_{zz}] = [K_1] + [A_1]'[H_1] = \begin{bmatrix} 65.995681 & 342.03116 & 1869.1504 \\ 342.03116 & 2039.9531 & 14673.089 \\ 1869.1504 & 14673.089 & 151545.20 \end{bmatrix}$$

$$[S^{22}{}_{zz}] = [K_2] + [A_2]'[H_2] = \begin{bmatrix} 75.556641 & 386.92007 & 1995.4171 \\ 386.92007 & 2243.7140 & 15044.211 \\ 1995.4171 & 15044.211 & 147808.27 \end{bmatrix}$$

and $$[S^{gg}{}_{zz}] = [K_g] + [A_g]'[H_g] = \begin{bmatrix} 8.2793523 & 38.357716 & 219.65978 \\ 38.357716 & 209.87519 & 1597.4279 \\ 219.65978 & 1597.4279 & 17693.325 \end{bmatrix}$$

Consequently, the SPR estimator of μ_1, μ_2, and μ_g are calculated as

$w_1 = [b]'[z^1] = 108150.31$ pounds

$w_2 = [b]'[z^2] = 120738.64$ pounds, and

$w_g = [b]'[z^g] = 12588.336$ pounds $= w_2 - w_1$

and the variances of w_1, w_2, and w_g as

$S_{w1w1} = [b]'[S^{11}{}_{zz}][b] + [z^1]'[S_{bb}][z^1] = 17292284 + 11049511 = 28341795$

$S_{w2w2} = [b]'[S^{22}{}_{zz}][b] + [z^2]'[S_{bb}][z^2] = 16553974 + 11720724 = 28274698$

and $S_{wgwg} = [b]'[S^{gg}{}_{zz}][b] + [z^g]'[Sbb][z^g] = 2078026.8 + 43968.8 = 2121995.6$

The reader can verify that

(1) The relative size of the second error component (due to the biomass regression function) is about 40 percent for the estimates of the current mean biomass/acre (μ_1 and μ_2) but about 2 percent for the estimate of the change. This is about twice as much as those we have found for the current values in the first example of the CFI without SPR and about the same for the change.

(2) In absolute values, the size of the first error components are very close to those obtained by the classical SPR estimators; the difference is less than one percent. The reason for the discrepancy is the fact that they are different estimators.

CONCLUDING REMARKS

We have shown an approach to combine the error of the sample plots with that of the biomass regression functions when biomass estimates are calculated in successive forest inventories with SPR. This approach requires that the biomass regression functions be of a linear form with good estimates of the (i) vector $[\beta]$ of regression coefficients and (ii) covariance matrix $[\sigma_{bb}]$ of the estimator [b] of $[\beta]$. The approach can be used with fixed area sample plots or relascope sample points. The sample plots are assumed to have been selected by simple random sampling.

An illustrative numerical example shows that the relative size of the error component due to biomass regression functions may sometimes be small (for the estimate of change between measurement times) but most of the time it may be sufficiently large so as not to be ignored, if a realistic estimate of the reliability of the inferences made is desired. In some instances, like the estimation of ingrowth, the size of the error due to biomass regressions is much higher than that of the sample plots.

REFERENCES

Cunia, T. (1986): On the error of biomass estimates in forest inventories: Part 2: The error component from sample plots. Faculty of Forestry Miscellaneous Publication Number 9, State University of New York, College of Environmental Science and Forestry, Syracuse NY.

Cunia, T. (1987): On the error of estimates from successive forest inventories with SPR. Canadian Journal of Forest Research, Vol. 17, No. 5, p. 442-447.

Cunia, T. (1989 a): Successive forest inventories: Basic SPR approach for two measurements. Proceedings of the "State of the Art Methodology of Forest Inventory" Symposium, July 31-August 6, 1989, SUNY College of Environmental Science and Forestry, Syracuse NY.

Cunia, T.(1989 b) : Successive forest inventories: Basic SPR approach for two measurements. Proceedings of the "State of the Art Methodology of Forest Inventory" Symposium, July 31-August 6, 1989, SUNY College of Environmental Science and Forestry, Syracuse NY.

Cunia, T. (1989 c): Successive forest inventories: Error of biomass regressions. Part I: CFI with permanent plots only. Proceedings of the "State of the Art Methodology of Forest Inventory" Symposium, July 31-August 6, 1989, SUNY College of Environmental Science and Forestry, Syracuse NY.

☆U.S. GOVERNMENT PRINTING OFFICE: 1990 -791-171/20042

LaBau, Vernon J.; Cunia, Tiberius, tech. eds. 1990. State-of-the-art methodology of forest inventory: a symposium proceedings; 1989 July 30-August 5; Syracuse, NY. Gen. Tech. Rep. PNW-GTR-263. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwst Research Station. 592 p.

A collection of about 85 papers presented at the Symposium on the State-of-the-Art Methodology of Forest Inventory, held at Syracuse University, Syracuse, New York, July 30-August 5, 1989. Includes workshop papers also presented as part of the symposium.

Keywords: Inventory, mensuration, biometrics, forest survey, remote sensing, proceedings symposium.

The **Forest Service** of the U.S. Department of Agriculture is dedicated to the principle of multiple use management of the Nation's forest resources for sustained yields of wood, water, forage, wildlife, and recreation. Through forestry research, cooperation with the States and private forest owners, and management of the National Forests and National Grasslands, it strives—as directed by Congress—to provide increasingly greater service to a growing Nation.

The U.S. Department of Agriculture is an Equal Opportunity Employer. Applicants for all Department programs will be given equal consideration without regard to age, race, color, sex, religion, or national origin.

Pacific Northwest Research Station
319 S.W. Pine St.
P.O. Box 3890
Portland, Oregon 97208-3890

U.S. Department of Agriculture
Pacific Northwest Research Station
319 S.W. Pine Street
P.O. Box 3890
Portland, Oregon 97208

Official Business
Penalty for Private Use, $300

BULK RATE
POSTAGE +
FEES PAID
USDA-FS
PERMIT No. G-40

do NOT detach label

75051
102

Clemson University
3 1604 009 855 307

DATE DUE

MAY 2 0 1994

Ret'd APR 07 1994

DEMCO, INC. 38-2931

INFORMATION ICI CONSERVATION, INC.